

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE ENGENHARIA GEOGRÁFICA, GEOFÍSICA E ENERGIA



Análise multitemporal de imagens de satélite para cartografia de produtos hortícolas

Ana Catarina Courela Ricardo

Mestrado em Engenharia Geográfica

Trabalho de Projeto orientado por:
Prof. Doutor João Catalão Fernandes
Prof.^a Doutora Ana Navarro

“Tenho a certeza de que hoje somos senhores do nosso destino, que a tarefa que temos perante nós não está acima das nossas forças e que as dores e dificuldades não estão para lá das nossas capacidades de resistência física. Enquanto tivermos fé na nossa própria causa e uma indominável força de ganhar, a vitória não nos será negada.”

Winston Churchill

AGRADECIMENTOS

Em jeito de conclusão deste trabalho de projeto bem como do percurso académico, é-me agora pertinente agradecer a todos aqueles que direta ou indiretamente estiveram presentes.

Primeiramente, tendo a consciência de que sozinha não teria sido fácil, obrigada aos meus pais por tudo. À minha mãe por ser um modelo de coragem e determinação e ao meu pai por ser o incentivo constante que é. Ao meu irmão Tiago, obrigada pela companhia e amizade.

À Professora Doutora Ana Navarro e ao Professor Doutor João Catalão, pela orientação, apoio, disponibilidade, opiniões críticas e total colaboração no solucionar de dúvidas e problemas que foram surgindo ao longo da realização deste trabalho. Impecáveis.

À PrimoHorto Sociedade de Produtores de Hortícolas, LDA e ao Instituto de Financiamento da Agricultura e Pescas, I. P. pela cedência de alguns dos dados utilizados na realização deste projeto.

À Professora Doutora Cristina Catita, por me ter dado a oportunidade de ser monitora de SIG, experiência esta que nunca terei suficientemente como agradecer.

Aos meus colegas e amigos que a faculdade trouxe, imprescindíveis ao bom rumo do curso, Susana Feio, Maria João Martins, Nuno Fragoeiro, Catarina Rodrigues, Inês Vilas Boas, Carolina Rocha e Joana Contente.

Às minhas amigas de infância e de sempre Cláudia Soares, Catarina Luz, Ana Rita Martins e às extraordinárias e aventureiras do BJJ, Inês Alves, Elisa de La Cruz, Leonor Lopes e Teresa Mendes Coelho.

Por fim, ao meu querido Ângelo, pela muita paciência, motivação e por fazer de mim uma pessoa mais curiosa.

Obrigada!

RESUMO

O presente estudo tem como finalidade desenvolver uma metodologia que permita identificar e cartografar parcelas hortícolas na zona agrícola do Montijo, com recurso a imagens de satélite da missão Sentinel-2. Para identificar as diferentes culturas presentes na zona em estudo, utilizou-se uma série temporal de 35 imagens dos satélites Sentinel-2A e Sentinel-2B abrangendo a época de crescimento das culturas, compreendida entre junho de 2017 e maio de 2018. A informação contida nas imagens permitiu, numa primeira fase, o cálculo do Índice de Vegetação de Diferença Normalizada (NDVI – *Normalized Difference Vegetation Index*) e o consequente estudo do ciclo de crescimento das culturas hortícolas ao longo do período temporal analisado. Posteriormente, foi desenvolvido um algoritmo que calcula o valor mínimo, máximo, médio e de desvio padrão do NDVI para cada píxel ao longo da série temporal, bem como a amplitude e o número de dias com base na curva do ciclo fenológico desde o início do desenvolvimento da cultura até à colheita. Este algoritmo permite gerar 6 bandas artificiais, as quais incorporam a variável temporal de cada uma das culturas hortícolas. Para o efeito foram testadas diferentes abordagens com vista à obtenção dos valores de amplitude e do número de dias mais consistentes com os que resultam da análise visual dos gráficos gerados com a série temporal de NDVI. Por fim, as 35 imagens de NDVI e as 6 bandas artificiais foram utilizadas separadamente para testar dois algoritmos de classificação supervisionada, nomeadamente o algoritmo da Máxima Verosimilhança (MV) e as Redes Neurais Artificiais (RNA). O algoritmo que apresentou melhores resultados foi o da MV, com uma exatidão global de 67.3%, sendo que a exatidão do produtor obtida para as culturas agrícolas analisadas foi bastante significativa (cerca de 99%). A metodologia desenvolvida proporcionou resultados que, estando naturalmente sujeitos a melhorias e servindo apenas como contributo científico, poderão futuramente ser utilizados para a verificação “remota” de declarações por parte dos agricultores, reduzindo consideravelmente o número de deslocações ao terreno e os consequentes custos que tal acarreta.

PALAVRAS CHAVE: Detecção Remota, Sentinel-2, NDVI, Culturas Hortícolas, Classificação Multitemporal, Máxima Verosimilhança, Redes Neurais Artificiais, *machine learning*.

ABSTRACT

The present study aims to develop a methodology that identifies and maps horticultural crops in the agricultural area of Montijo, using satellite images from the Sentinel-2 mission. To identify the different types of cultures present in this study area, we used a time series of 35 images from the Sentinel-2A and Sentinel-2B satellites, that covers the growing season between June 2017 and May 2018. The information contained in the images allowed, in the first stage, the calculation of the Normalized Difference Vegetation Index (NDVI) and the consequent study of the horticultural crops growth cycle over the analyzed time period. Then, an algorithm was developed to calculate the minimum, maximum, average and standard deviation of the NDVI for each pixel along the time series, as well as the amplitude and the number of days based on the phenological cycle curve since the beginning of crop development until harvesting. This algorithm enables the calculation of 6 artificial bands which incorporate temporal variable for each horticultural crop. For this purpose, different approaches were tested in order to obtain the most consistent values of amplitude and number of days with those that were retrieve from the visual analysis of the NDVI time series graphs. Finally, the 35 NDVI images and the 6 artificial bands were used separately to evaluate two supervised classification algorithms, namely the Maximum Likelihood (ML) and the Artificial Neural Networks (ANN). The algorithm that produced the best results was the ML, with an overall accuracy of 67.3%, and a very significant producer's accuracy for the horticulture crops (around 99%). The developed methodology yielded results that, although naturally subject to improvement and considered as just a scientific contribution, might be used in the future to validate remotely farmers declarations, and to considerably reduce the number of field trips and the consequently the costs involved.

KEYWORDS: Remote Sensing, Sentinel-2, NDVI, Horticultural Crops, Multitemporal Classification, Maximum Likelihood, Artificial Neural Network, machine learning.

ÍNDICE

Agradecimentos	iii
Resumo	iv
Abstract	v
Índice	vi
Lista de Tabelas	viii
Lista de Figuras	ix
Acrónimos	x
Capítulo 1	1
1. Introdução.....	1
1.1. Enquadramento.....	1
1.2. Motivação e Objetivos.....	3
1.3. Estrutura do Trabalho de Projeto.....	3
Capítulo 2	4
2. Estado da Arte	4
Capítulo 3	8
3. Dados e Métodos	8
3.1. Caracterização da Área de Estudo	8
3.2. Dados Utilizados	9
3.2.1. Imagens de Satélite.....	10
3.2.2. Informação das Parcelas Hortícolas	11
3.2.3. Sistema de Identificação de Parcelas.....	12
3.3. Métodos	13
3.3.1. Cálculo do Índice de Vegetação NDVI.....	14
3.3.2. Avaliação Fenológica: criação e estudo de Curvas de Crescimento por Parcela	14
3.3.3. Desenvolvimento do Algoritmo	15
3.3.4. Criação de uma Máscara	16
3.3.5. Classificação e Validação.....	17
3.3.5.1. Máxima Verosimilhança (MV)	17
3.3.5.2. Redes Neurais Artificiais (RNA).....	18
3.3.5.3. Análise da Qualidade da Classificação.....	19
Capítulo 4	20
4. Resultados e Discussão	20
4.1. Avaliação da Fenologia das Culturas	20
4.1.1. Análise da Evolução Temporal do NDVI para culturas hortícolas	20
4.1.2. Evolução Temporal do NDVI para culturas do IFAP	22
4.2. Algoritmo para determinação dos Parâmetros característicos da Curva de NDVI.....	23
4.2.1. Bandas Artificiais geradas	23
4.2.2. Máscara com Zonas de Ocupação Hortícola	26
4.2.2.1. Composição Colorida.....	27
4.3. Classificação Segundo os Métodos de Máxima Verosimilhança (MV) e Redes Neurais Artificiais (RNA).....	29

4.3.1.	Comparação entre MV e RNA	29
4.3.2.	Estudo das RNA com máscara	33
4.3.2.1.	Parametrização de modelos de RNA	33
Capítulo 5	37
5.	Conclusões e Perspetivas Futuras.....	37
5.1.	Principais Conclusões.....	37
5.2.	Perspetivas Futuras.....	38
Referências bibliográficas	39
Anexos	43
A. Datas de sementeira e colheita das Parcelas seleccionadas para o período de junho a dezembro de 2017, fornecidas pela PrimoHorto	43
B. Evolução temporal do NDVI por parcelas (junho 2017 - maio 2018)	45
i.	Parcelas de Couve	45
ii.	Parcelas de Batata.....	45
iii.	Parcelas de Cenoura	47
C. Valores médios por parcela calculados mediante análise visual em Excel	49
i.	Parcelas de Couve	49
ii.	Parcelas de Batata.....	49
iii.	Parcelas de Cenoura	49
D. Comparação do número de dias calculados entre as versões dos algoritmos	50
E. Valores médios por parcela calculados através da Versão 1 do algoritmo em Matlab	51
i.	Parcelas de Couve	51
ii.	Parcelas de Batata.....	51
iii.	Parcela de Cenoura.....	51
F. Máscara aplicada	52
G. Dados de treino e de validação para a comparação de resultados na classificação das culturas usando MV e RNA	52
i.	Dados de Treino	52
ii.	Dados de Validação.....	53
H. Dados de treino e de validação para a comparação de resultados na classificação das culturas mediante os diferentes parâmetros das RNA	53
i.	Dados de Treino	53
ii.	Dados de Validação.....	54

LISTA DE TABELAS

Tabela 3.1 – Lista das aplicações informáticas utilizadas e respetiva função.....	9
Tabela 3.2 – Data de aquisição das imagens Sentinel-2A e Sentinel-2B, sob correção atmosférica Nível-1C.....	10
Tabela 3.3 – Principais características das imagens Sentinel-2.....	10
Tabela 3.4 – Percentagem de ocupação de cada cultura em estudo.	11
Tabela 3.5 – Culturas controladas para dados de treino e validação.....	12
Tabela 3.6 – Valores médios de NDVI por classe (média das médias e média do desvio padrão).....	17
Tabela 3.7 – Desempenho da classificação segundo os valores do índice Kappa [adaptado de Landis e Koch, 1977].....	19
Tabela 4.1 – Valores Médios calculados para o ciclo fenológico das 16 parcelas de Batata em Excel.	20
Tabela 4.2 – Valores Médios calculados para o ciclo fenológico das 16 parcelas de Cenoura em Excel.	21
Tabela 4.3 – Valores Médios calculados para o ciclo fenológico das 2 parcelas de Couve em Excel..	22
Tabela 4.4 – Valores médios calculados por classe de parcela através da versão 1 do algoritmo em MATLAB®.....	24
Tabela 4.5 - Número de Parcelas de Treino (T) e de Validação (V) consideradas na comparação entre MV e RNA.	29
Tabela 4.6 – Resultados da exatidão global e do índice Kappa na comparação entre MV e RNA.....	29
Tabela 4.7 – Resultados estatísticos para couve, cenoura e batata aquando o uso das 35 bandas de NDVI.....	30
Tabela 4.8 – Resultados estatísticos para couve, cenoura e batata aquando o uso das 6 bandas artificiais de parâmetros.	31
Tabela 4.9 – Exatidões do Produtor e do Utilizador para píxeis de batata e cenoura juntos (b+c) para as 35 bandas de NDVI e as 6 bandas artificiais.....	31
Tabela 4.10 – Comparação visual dos resultados das classificações das 35 e 6 bandas (EG – Exatidão Global, IK – Índice Kappa, EP – Exatidão do Produtor e EU – Exatidão do Utilizador).....	32
Tabela 4.11 – Número de Parcelas de Treino (T) e de Validação (V) consideradas na comparação de Parâmetros para RNA.....	33
Tabela 4.12 – Parâmetros considerados nas 20 configurações das RNA e respetivas estatísticas de classificação.....	33
Tabela 4.13 – Melhores resultados estatísticos para couve, cenoura e batata aquando comparação de parâmetros em RNA.....	34
Tabela 4.14 – Exatidões do Produtor e do Utilizador para píxeis de batata e cenoura juntos (b+c) para 6 bandas com máscara.....	35
Tabela 4.15 – Comparação visual dos resultados das classificações com diferentes parâmetros de entrada nas RNA (EG – Exatidão Global, IK – Índice Kappa, EP – Exatidão do Produtor e EU – Exatidão do Utilizador).....	36

LISTA DE FIGURAS

Figura 3.1 – Subset aplicado na Área de Estudo.	8
Figura 3.2 – Valores de precipitação média e temperatura média das normais climatológicas entre 1981 e 2010 para a estação de Setúbal [adaptado de IPMA, 2018].	9
Figura 3.3 – Parcelas Hortícolas usadas no estudo da fenologia das culturas de batata, cenoura e couve.	11
Figura 3.4 – Parcelas Hortícolas usadas para teste e validação da classificação.	12
Figura 3.5 – Diagrama do trabalho de projeto.	13
Figura 3.6 – Representação gráfica do ciclo de crescimento teórico das culturas hortícolas com os parâmetros considerados.	15
Figura 3.7 – Representação esquemática do algoritmo aplicado na criação das 6 bandas artificiais. ...	15
Figura 3.8 – Ilustração do comportamento da Versão 1, Versão 2 e Versão Half na escolha de máximos e mínimos para um mesmo píxel da série temporal, sendo (b) função inversa de (a). ..	16
Figura 3.9 – Rede Neuronal de multicamadas. Adaptado de Richards e Jia [2006].	18
Figura 4.1 – (a) Evolução temporal do NDVI de todas as parcelas de batata estudadas. (b) Evolução temporal do NDVI para uma parcela genérica de batata (05-Quinta da Vinha).	21
Figura 4.2 – (a) Evolução temporal do NDVI de todas as parcelas de cenoura estudadas. (b) Evolução temporal do NDVI para uma parcela genérica de cenoura (09B-Valmet).	21
Figura 4.3 – (a) Evolução temporal do NDVI de todas as parcelas de couve estudadas. (b) Evolução temporal do NDVI para uma parcela genérica de couve (16-Rotunda).	22
Figura 4.4 – Tendência média do NDVI para cada tipo de cultura.	22
Figura 4.5 – Ciclo Fenológico das culturas de validação fornecidas pelo IFAP.	23
Figura 4.6 – Banda artificial representativa do valor médio e NDVI para cada píxel ao longo da série temporal, calculado através da função mean do MATLAB®, embutida na Versão 1 do Algoritmo.	24
Figura 4.7 – Banda artificial representativa do desvio padrão de NDVI para cada píxel ao longo da série temporal, calculado através da função std do MATLAB®, embutida na Versão 1 do Algoritmo.	25
Figura 4.8 – Banda artificial representativa do valor mínimo de NDVI para cada píxel ao longo da série temporal, calculado através da função min do MATLAB®, embutida na Versão 1 do Algoritmo.	25
Figura 4.9 – Banda artificial representativa do valor máximo de NDVI para cada píxel ao longo da série temporal, calculado através da função max do MATLAB®, embutida na Versão 1 do Algoritmo.	25
Figura 4.10 – Banda artificial representativa do número de dias para cada píxel ao longo da série temporal, calculados através da Versão 1 do algoritmo em MATLAB®.	26
Figura 4.11 – Banda artificial representativa da amplitude da curva para cada píxel ao longo da série temporal, calculados através da Versão 1 do algoritmo em MATLAB®.	26
Figura 4.12 – Valores médios da Média e Desvio Padrão de NDVI, por classes, ao longo da série temporal.	27
Figura 4.13 – Composição colorida das bandas número de dias (R), desvio padrão (G) e média (B) após a aplicação da máscara binária.	27
Figura 4.14 – Detalhe da composição colorida com algumas das parcelas em estudo.	28

ACRÓNIMOS

BSQ	Band Sequencial
CCRS	Canada Centre for Remote Sensing
DR	Deteção Remota
ESA	European Space Agency
IFAP	Instituto de Financiamento da Agricultura e Pescas
MV	Máxima Verosimilhança (<i>Maximum Likelihood</i>)
NaN	<i>Not a Number</i>
NDVI	<i>Normalized Difference Vegetation Index</i> (Índice de Vegetação de Diferença Normalizada)
NIR	<i>Near Infrared</i> (Infravermelho Próximo)
NUTS	Nomenclatura das Unidades Territoriais para Fins Estatísticos
RNA	Redes Neurais Artificiais (<i>Artificial Neural Network</i>)
S2	Missão Sentinel-2
SIG	Sistemas de Informação Geográfica
SWIR	<i>Short-Wave Infrared</i> (Infravermelho de Onda Curta)
TOA	Top Of Atmosphere
USGS	United States Geological Survey
WGS84/UTM	World Geodetic System 1984/ Universal Transversa de Mercator

CAPÍTULO 1

INTRODUÇÃO

1.1. ENQUADRAMENTO

“Segundo estimativas da Food and Agriculture Organization das Nações Unidas, os agricultores terão que, até 2050, aumentar a sua produção em 70%, para que seja possível alimentar as cerca de 9,1 mil milhões de pessoas que habitarão no nosso planeta, por essa altura. O cumprimento destes objetivos depende da otimização e eficiência das culturas, que poderá ser potenciada pela conectividade entre equipamentos e por sistemas baseados em sensores, permitindo controlar todo o processo de forma minuciosa e em tempo real, desde o planeamento até à entrega do produto ao consumidor, seja na zona de cultivo ou além-fronteiras [FAO, 2009].”

Nos tempos antecedentes à Revolução Industrial, a Agricultura evidenciava-se como uma prática de subsistência, de pequena escala e com recurso a métodos de produção muito artesanais e manuais. A falta de equipamentos agrícolas levava a que, mesmo numa pequena parcela familiar, a mão-de obra fosse insuficiente. A partir da Revolução Industrial (final do século XVIII e início do século XIX), com o desenvolvimento das técnicas até então aplicadas, o Homem passa a ter acesso à maquinaria agrícola, permitindo o aumento de produção em paralelo com a redução de custos e de mão-de-obra. O principal fator de estímulo ao desenvolvimento e à modernização da agricultura foi a acumulação de bens de capital, que proporcionaram um aumento da capacidade de financiar máquinas modernas e, assim, o consequente aumento da produtividade agrícola.

O que a história refere como a grande impulsão na agricultura desde a Revolução Industrial, começa a alterar-se nos dias de hoje, em grande parte devido à consciência ambiental. A introdução de fitoquímicos e fertilizantes de síntese começou a ser vista como um ponto a alterar, ou pelos menos a sua forma de aplicação. Ao contrário do que se fazia sentir antes da utilização destes fatores de produção, hoje os campos agrícolas são tratados de forma homogénea, obtendo-se no fim da campanha uma produção heterogénea (em quantidade e qualidade), devido ao desprezo dado a outros fatores que geram grande variabilidade para a cultura, como o solo e a topografia. A necessidade de se alterar este rumo originou a criação uma outra forma de encarar a produção e gestão agrícola, surgindo o conceito de agricultura de precisão no final do século XX. Este conceito está normalmente associado ao uso de equipamentos de alta tecnologia, para avaliação e monitorização das condições de uma parcela de terreno, aplicando-se depois os diversos fatores de produção em conformidade [Neto *et al.*, 2005].

A Deteção Remota (DR) é usada na agricultura de precisão para detetar a variabilidade espacial e temporal nas condições do solo e das culturas. As imagens de satélites são utilizadas há muito tempo pelos mais diversos setores, porém, foi através dos avanços tecnológicos e da redução significativa dos custos que se tornou viável a aplicação em larga escala destas tecnologias na agricultura. A possibilidade de controlar as culturas e tomar medidas à distância, permite a otimização de processos e a utilização de recursos bem como a gestão do território de forma mais eficiente. A digitalização da agricultura nesta era moderna, permite não só um novo incremento da produção e qualidade bem como a diminuição dos riscos e custos associados.

No caso da DR por satélite multiplicam-se as plataformas de acesso gratuito quer a imagens RGB, de falsa-cor e a outros compósitos mais específicos para a agricultura quer a índices de vegetação, tais como o (re)conhecido Índice de Vegetação de Diferença Normalizada (NDVI – *Normalized Difference Vegetation Index*). Algumas das plataformas apresentam imagens quer do Sentinel-2 quer ainda do LANDSAT ou MODIS. A disponibilidade gratuita destas imagens de considerável resolução espacial e temporal, como é atualmente o caso da missão Sentinel-2, abre assim novas possibilidades para a análise e monitorização dos recursos agrícolas. Recorrendo a séries temporais de imagens de satélites e a Sistemas de Informação Geográfica (SIG), é possível avaliar a dinâmica das alterações do uso e ocupação do solo e por conseguinte, contribuir com ferramentas de suporte à gestão e ao planeamento de culturas hortícolas.

Embora a discriminação de culturas seja uma etapa crítica para a maioria dos sistemas de monitorização agrícola, a capacidade de identificar o tipo de cultura torna-se imprescindível na estimação da área associada a cada parcela, permitindo obter variáveis estatísticas essenciais ao controle das culturas, nomeadamente em áreas subsidiadas, e ao fornecimento de dados de entrada para modelos de previsão de produção agrícola [Zilman e Weichelt, 2014]. O problema da classificação de culturas pode ser resolvido mediante diferentes abordagens, onde a identificação/discriminação das mesmas passa pela utilização de algoritmos de classificação que não se baseiem apenas na informação ao nível do píxel, podendo ser baseados em objetos, assim como em técnicas híbridas que combinam dois ou mais métodos. Existem algoritmos mais sofisticados e igualmente complexos, como os de *machine learning*, onde, tal como o nome da indica, são algoritmos que aprendem com os dados de treino e à medida que a curva de aprendizagem dos mesmos aumenta, os seus resultados melhoram.

Hoje em dia são várias as técnicas para monitorizar a vegetação e detetar alterações, contudo, na prática, a seleção do método mais adequado não é fácil, sendo o objetivo da análise e a área de estudo fatores decisivos na sua escolha.

1.2. MOTIVAÇÃO E OBJETIVOS

A classificação multiespectral de dados em DR tem sido amplamente utilizada na criação de inventários temáticos de uso e ocupação do solo, sob grande variedade de aplicações que incluem o planejamento urbano, a classificação de florestas e a caracterização de culturas agrícolas [Yuan *et al.*, 2009]. Com este trabalho de projeto pretende-se desenvolver uma metodologia passível de identificar e monitorizar culturas hortícolas presentes na área de estudo, recorrendo a uma série temporal de imagens óticas dos satélites Sentinel-2A e 2B.

Para o efeito é efetuada uma análise multitemporal com recurso a uma série de imagens de satélite adquiridas ao longo do ciclo de crescimento das culturas analisadas e a informação sobre os limites e o tipo de culturas das respetivas parcelas agrícolas. Para tal, é analisada a fenologia das culturas hortícolas através do cálculo do valor médio do NDVI para cada parcela e para cada uma das datas observadas de modo a caracterizar as distintas fases do ciclo de crescimento das culturas e a identificar a sua duração (período vegetativo).

Será ainda desenvolvido um algoritmo que calcula parâmetros com base na informação da série temporal de imagens de NDVI (píxel a píxel), gerando 6 bandas artificiais para efeitos de classificação. Adicionalmente, pretende-se estudar o desempenho das Redes Neurais Artificiais em comparação com o do tradicional classificador da Máxima Verosimilhança.

1.3. ESTRUTURA DO TRABALHO DE PROJETO

O presente documento encontra-se estruturado em cinco capítulos, divididos em subcapítulos. No primeiro capítulo é efetuado um enquadramento do tema do trabalho realizado e são apresentados os objetivos a cumprir. O segundo capítulo destina-se ao estado da arte, referindo a realidade do conhecimento atual e contextualizando a metodologia implementada com literatura revista no âmbito do tema. No capítulo seguinte é feita a caracterização da área de estudo e são apresentados os dados e a metodologia utilizada para a realização do trabalho. No capítulo quatro são apresentados os resultados obtidos e é realizada uma análise aos mesmos. Por último, no capítulo cinco, são tecidas algumas conclusões relativas aos resultados obtidos e apresentadas algumas perspetivas de desenvolvimentos futuros.

CAPÍTULO 2

ESTADO DA ARTE

O mapeamento do uso e cobertura do solo é crucial para compreender, monitorizar e prever os efeitos da complexa interação Homem x Natureza nas escalas locais, regionais e globais [Clark *et al.*, 2010]. No entanto, usar apenas inventários de campo para mapear extensas áreas de vegetação pode tornar-se numa tarefa morosa e significativamente cara, o que leva ao uso cada vez mais frequente de imagens aéreas e de satélite para atender aos mesmos propósitos [Bradter *et al.*, 2011]. As tecnologias de DR têm provado ser uma eficiente abordagem na documentação, descrição e enumeração da cobertura de uso e ocupação do solo, sendo um fator importante na formação de políticas relativas às preocupações económicas, demográficas e ambientais [Mishra *et al.*, 2017].

As culturas hortícolas, constituídas por legumes, frutas, flores, especiarias, plantações e plantas medicinais desempenham um papel significativo na melhoria da produtividade da terra, gerando emprego, aumentando as exportações, melhorando as condições económicas dos agricultores e empresários e fornecendo segurança alimentar e nutricional para as pessoas. Para uma melhor gestão das culturas existentes e para definir melhor áreas de culturas hortícolas, é necessária uma base de dados atualizada e precisa para o planeamento sistemático e a tomada de decisões. A DR como ferramenta avançada que é, auxilia a recolha e atualização de informações no desenvolvimento de planos de gestão, podendo ser usada na redução de riscos e na minimização de danos [Usha e Singh, 2013]. Os mesmos dados podem também ser analisados de maneiras diferentes em diferentes aplicações. Vários estudos visam por exemplo a identificação de culturas, a estimativa de áreas, a monitorização de incêndios ou cheias, ou a identificação de doenças e pragas, usando dados de satélite [Karkee *et al.*, 2009; Hebbar *et al.*, 2014; Petropoulos *et al.*, 2010; Acharya e Thapa, 2015].

Muitos tipos de sensores, como radiómetros de microondas, medidores de laser, sensores magnéticos e câmaras, recolhem informações eletromagnéticas para obter informações precisas e em larga escala sobre a superfície e a atmosfera da Terra. Como os dados e as imagens são digitais, podem facilmente ser quantificados e manipulados usando computadores. Os sistemas de DR, devido ao uso regular, sinóptico, multiespectral e temporal de uma área, fornecem uma base de dados precisa sobre o comportamento espectral das culturas, bem como do seu ambiente de crescimento, ou seja, solo e atmosfera. Estes dados podem ser usados para uma série de aplicações, tais como a elaboração de inventários de colheita, a avaliação das condições de colheita, a elaboração de previsões de produção, a avaliação da qualidade da fruta, o cálculo do índice de área foliar, a deteção do crescimento e da saúde das culturas hortícolas, a avaliação de seca ou de danos de inundação, e a monitorização e gestão do solo [Min *et al.*, 2008; Mondal e Basu, 2009].

Para a identificação precisa do uso e ocupação do solo de áreas agrícolas a uma grande escala são necessárias imagens de DR com alta resolução espacial e temporal [Dusseux *et al.*, 2014]. As imagens obtidas pelos satélites Sentinel-2A e Sentinel-2B, geridos pela Agência Espacial Europeia e desenvolvidos especificamente para as necessidades operacionais do programa Copernicus, proporcionam uma melhoria significativa na cobertura espectral, na resolução espacial e na frequência de revisita comparativamente ao sistema de satélites Landsat. No total, os sensores têm 13 bandas

espectrais, listadas como banda 1-8, 8A e 9-12, com uma resolução espacial de 10 m (banda 2, 3, 4 e 8), 20 m (banda 5, 6, 7, 8A, 11 e 12) e 60 m (banda 1, 9 e 10), as quais fornecem dados para aplicações de média e alta resolução [ESA, 2018]. Os últimos anos e podem inclusivamente ser chamados de anos de *Big Free Data* em DR. Durante o período 2013–2017 foram lançados vários satélites de radar de abertura ótica e sintética (SAR), igualmente de elevada resolução espacial e os Landsat-8 dentro do Projeto Landsat, uma iniciativa conjunta entre o US Geological Survey (USGS) e a National Aeronautics and Space Administration [Roy *et al.*, 2014]. A entrada em órbita do 2º satélite da missão Sentinel-2 a 7 de março de 2017 oferece uma combinação sem precedentes da cobertura global e sistemática da superfície terrestre, onde o amplo campo de visão para observações multiespectrais e o período de revisita de cinco dias permitem a monitorização de fenómenos de evolução rápida, tal como o desenvolvimento de culturas agrícolas.

As características espectrais do coberto vegetal existente na superfície terrestre variam consideravelmente no tempo e no espaço. A refletância espectral da vegetação apresenta uma grande variação em função das estações do ano, da espécie vegetal, do teor de água na planta e do comprimento de onda do sensor [Fonseca e Fernandes, 2004]. A clorofila (composto químico presente nas folhas) é um fator condicionante da resposta espectral, pois esta absorve fortemente os comprimentos de onda do vermelho e azul. Quando o teor de clorofila da vegetação está no máximo existe uma maior absorção dos comprimentos de onda vermelhos e maior reflexão de infravermelho. Por outro lado, quando há menos clorofila nas folhas, existe uma menor absorção e proporcionalmente uma maior reflexão dos comprimentos de onda vermelhos [CCRS, 2018]. Os índices de vegetação baseiam-se nas diferenças de refletividade que a vegetação de cor verde apresenta nas regiões do visível e do infravermelho. Têm dois objetivos principais, realçar aspetos que permitem discriminar entre diferentes tipos e estados de vegetação e minimizar radiâncias do mesmo tipo de coberto, resultantes de diferenças de iluminação devidas à topografia [Fonseca e Fernandes, 2004].

Os índices de vegetação fornecem informação sobre as condições da vegetação através da relação entre as bandas do satélite e as características das culturas. Embora existam muitos índices de vegetação, o NDVI é o mais utilizado a nível global. Desenvolvido por Rouse *et al.* [1973], o NDVI é calculado a partir das refletâncias espectrais ao nível do topo da atmosfera das bandas do vermelho (*Red*) e do infravermelho próximo (*NIR*), sendo obtido através da divisão da diferença de valores da refletância do infravermelho próximo e do vermelho pela soma entre os mesmos valores de refletância. Estes comprimentos de onda são os mais sensíveis à absorção, reflexão e transmissão de energia da clorofila presente e também à densidade de vegetação verde na superfície do solo. A junção destes dois comprimentos de onda (zona de baixa refletância com uma zona de alta refletância) permite obter o contraste máximo entre o solo e a vegetação [Warren e Metternicht, 2005].

A dinâmica das culturas produz variações de longo e curto prazo nas suas propriedades no decorrer do tempo. Essas variações são reguladas por fatores ambientais, tais como a temperatura, a radiação, a humidade, a quantidade de nutrientes, entre outros. A monitorização da dinâmica das culturas é essencial para observar as mudanças e tendências nesses fatores e consequências subsequentes para os recursos da vegetação e uso humano deles. A fenologia, do grego *phainein*, estuda as mudanças externas visíveis nas plantas durante o seu desenvolvimento vegetativo e reprodutivo, onde a resposta de cada planta face às variações sazonais e aos fatores ambientais se manifesta na floração, no surgimento de folhas ou na queda das mesmas, por exemplo [Menzel, 2002]. Existe um grande corpo de literatura científica sobre o uso de índices de vegetação, como o NDVI, para a monitorização a longo prazo do crescimento da vegetação em escalas regionais e globais, e em diferentes zonas climáticas [Piao *et al.*, 2003; Huang *et al.*, 2014]. Também a fenologia e as suas alterações foram extensivamente estudadas [Heumann *et al.*, 2007; Jönsson *et al.*, 2010].

Alguns estudos recentes exploraram a dinâmica da vegetação usando dados de alta resolução, como os dos sensores Landsat e SPOT [Masek *et al.*, 2008; Fraser *et al.*, 2011; Tagesson *et al.*, 2012]. Embora certos aspectos da variação dinâmica possam ser monitorizados, a baixa resolução temporal impede o mapeamento preciso da fenologia ou distúrbios no perfil sazonal. Neste sentido, os satélites Sentinel-2 em conjunto permitem o mapeamento da fenologia e dinâmica sazonal com uma alta resolução espacial e temporal, superando assim muitos dos problemas relacionados com o uso de dados de baixa resolução. Eklundh *et al.* [2012] desenvolveram e testaram uma metodologia piloto para a geração de informação sobre a dinâmica da vegetação a partir de dados de alta resolução espacial, como os da missão Sentinel-2. Para tal, foi usada uma extensão do *software* TIMESAT que gerava perfis sazonais de NDVI suavizados e consequentes parâmetros de sazonalidade, como o início e o fim, o comprimento, a amplitude, os valores integrados, o máximo sazonal e derivativos, entre outros. O estudo demonstrou que o TIMESAT foi bem-sucedido em ajustar as funções do modelo na suavização dos dados e em gerar parâmetros de sazonalidade para a área de teste, para uma resolução espacial de 10 x 10 m.

O uso de séries temporais de imagens de satélites para classificar a vegetação tem sido um tópico de muitas pesquisas ao longo do tempo, particularmente à medida que as capacidades temporais dos sistemas de satélites têm vindo a melhorar de forma consistente. Diversos autores têm constatado que a utilização de imagens adquiridas em várias datas possibilita uma melhoria significativa na classificação de culturas [Karjalainen *et al.*, 2008; Long *et al.*, 2013], remetendo para a importância de análises multitemporais neste tipo de estudos e referindo que modelos multitemporais produzem melhores resultados que dados de uma única data, uma vez que estes registam a diversidade espectral devido às diferenças fenológicas. Também Zillmann e Weichelt [2014] referem que a classificação de culturas é um fator chave na monitorização agrícola, beneficiando da análise multitemporal de imagens correspondentes ao seu período de crescimento, utilizando o comportamento fenológico distinto de diferentes culturas. No entanto, a resposta espectral de algumas culturas pode ser semelhante em determinados períodos, mas pode também mostrar diferenças durante a época de crescimento, permitindo a sua discriminação. As limitações à identificação de culturas com base em imagens de satélite estão relacionadas com fatores tais como, a semelhança da refletância de plantas de diferentes culturas, a variação da refletância de uma mesma cultura em diferentes parcelas e as mudanças no padrão da fenologia de uma cultura [Yang *et al.*, 2011].

Definida por Singh [1989] a deteção de alterações (DA) é o processo de identificação de diferenças do estado de um objeto ou fenómeno através da sua observação em diferentes momentos, sendo útil em diversas aplicações como análise da ocupação do solo, monitorização agrícola, estudo fenológico da vegetação, mudanças sazonais, deteção de stress na vegetação, monitorização do degelo e outras mudanças ambientais. Para a classificação de culturas, as DA podem basear-se em objetos, sendo obtidos melhores resultados comparativamente aos obtidos com classificações ao nível do píxel. Existem diversas técnicas de DA, sendo a mais usual a aplicação de um valor limite (*threshold*) para distinguir entre mudança e não mudança. Os métodos mais simples consideram operações baseadas no píxel, enquanto que os mais complexos se baseiam em objetos. O primeiro passo da classificação de imagens consiste na elaboração de uma legenda apropriada com todas as classes de ocupação do solo necessárias ao estudo, contendo todas as possibilidades de tipos de ocupação [CCRS, 2018]. Existem vários métodos de classificação, que através de diversas abordagens procuram identificar com eficácia a informação de cada píxel ou objeto da imagem, classificando-o em categorias de acordo com a sua informação espectral.

O classificador de Máxima Verosimilhança (MV) é um dos algoritmos mais utilizados no processamento de imagens de satélite. Apesar de computacionalmente não ser dos mais eficientes, é um dos classificadores que apresenta melhores resultados. O algoritmo assume que as assinaturas espectrais das

áreas de treino de cada classe apresentam uma distribuição Gaussiana, e para isso, avalia a média, variâncias e covariâncias das áreas de treino, para assim determinar a distribuição Gaussiana de cada uma delas. É um classificador considerado paramétrico, pois envolve parâmetros, vetor da média e matriz de covariância, da distribuição gaussiana multivariada, sendo supervisionado, pois estima estes parâmetros através das amostras previamente definidas [Exelis, 2018]. Por sua vez, as Redes Neurais Artificiais (RNA) pertencem a técnicas de inteligência artificial, que são ferramentas computacionais amplamente utilizadas na análise de imagens. De acordo com Haykin [1994] uma RNA é um processador massivamente distribuído em paralelo, composto de unidades de processamento simples, que têm uma propensão natural para armazenar conhecimento experiencial e torná-lo disponível para uso. Uma RNA típica compreende um grande número de unidades de processamento simples, chamadas de nós, ligadas por conexões ponderadas de acordo com uma arquitetura especificada. O modelo básico das RNA consiste numa camada de entrada, uma camada oculta e uma camada de saída. A aprendizagem ocorre ajustando os pesos no nó para minimizar a diferença entre a ativação do nó de saída e a saída. O utilizador pode selecionar o número de camadas ocultas a serem usadas e escolher entre uma função de ativação logarítmica ou hiperbólica. Para a implementação da RNA, é necessário definir um número de parâmetros possíveis ao ajuste da rede [Petropoulos *et al.*, 2010]. Geralmente, é um algoritmo mais moroso, constituído por unidades de processamento que aplicam funções matemáticas capazes de reconhecer padrões e realizar atividades variadas aos dados recebidos.

O classificador estatístico convencional MV é mais frequentemente utilizado para classificações em DR, pois supõe-se que os dados de origem tenham uma distribuição normal. No entanto, os dados de várias fontes, especialmente dados auxiliares geográficos, podem não ter distribuição normal. Além disso, pode tornar-se mais difícil calcular a matriz de covariância requerida pela MV com dados de grande dimensão. Teoricamente este défice pode ser resolvido recorrendo a RNA, uma vez que os dados de entrada não têm necessariamente de ter uma distribuição normal [He *et al.*, 2006]. Embora o objetivo final seja a classificação temática, classificadores diferentes levam as classificações diferentes, sendo estes não só influenciados pela quantidade e qualidade dos dados de treino como também pela informação espectral disponível. O objetivo da análise e a área de estudo são, portanto, fatores decisivos na sua escolha de um classificador. Diversos estudos têm identificado o algoritmo MV e outros classificadores (*Support Vector Machine*, *Spectral Angle Mapper*) como melhores classificadores em resposta a melhores exatidões globais face às RNA [Mishra *et al.*, 2017; Nijhawan *et al.*, 2017], no entanto, o contrário também se verifica [Aziz *et al.*, 2017].

CAPÍTULO 3

DADOS E MÉTODOS

3.1. CARACTERIZAÇÃO DA ÁREA DE ESTUDO

A área de estudo, de aproximadamente 33000 ha, é definida pela zona abrangida pelos concelhos de Montijo, Alcochete, Palmela e Moita, pertencentes ao distrito de Setúbal, estando estes inseridos na NUTS III Península de Setúbal, e, por sua via, na NUTS II Lisboa (Figura 3.1).

A atividade agrícola na margem sul do Rio Tejo tem uma tradição de longa data, tendo a sua origem e expansão resultado da influência e proximidade de Lisboa, do seu porto e do seu mercado. A agricultura destes concelhos baseia-se na produção de culturas hortícolas, com destaque para a batata, cebola, cenoura, couves e hortícolas de estufa, sendo praticadas em áreas tão díspares que vão desde a horta familiar até ao cultivo extensivo em grandes áreas.

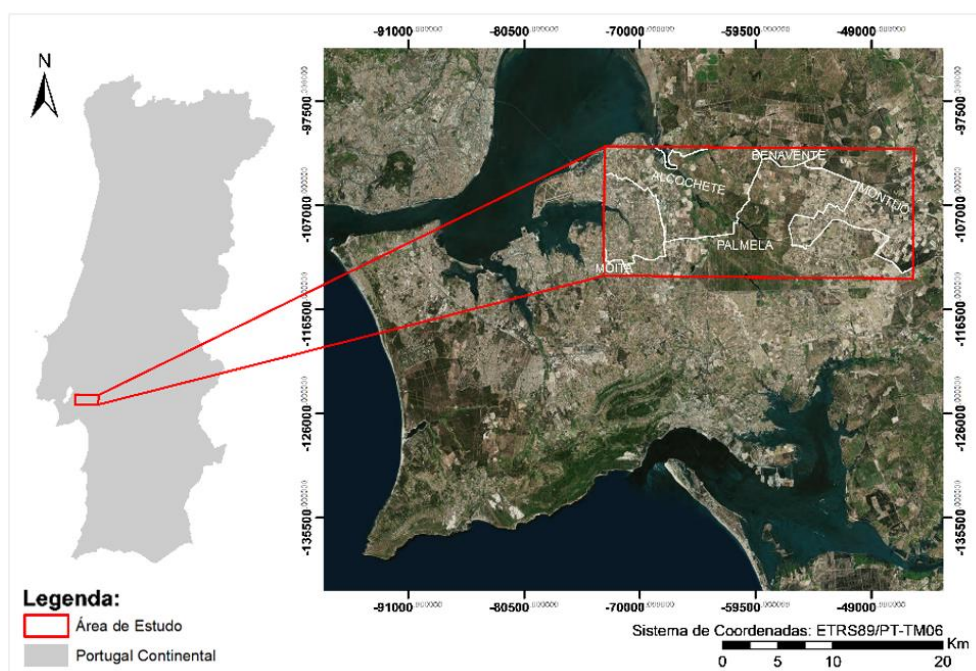


Figura 3.1 – Subset aplicado na Área de Estudo.

O clima da região, de influência mediterrânica é classificado pelo critério mais simples como temperado e moderado chuvoso, de poucas geadas e fraca amplitude térmica. Invernos frios e verões quentes. O clima desta região é condicionado sobretudo pela proximidade do oceano, embora seja atenuado a Sul pela serra da Arrábida, que forma uma espécie de “muralha” entre esse território e o mar.

O ar, cuja temperatura média anual é de 16.6°C, é característico de um clima temperado. Apenas em agosto sobe para uma média de 23°C e em janeiro, o mês mais frio, desce até 10.1°C [adaptado de IPMA, 2018]. As chuvas, bem como as maiores geadas chegam nos meses mais frios (dezembro, janeiro e fevereiro). A humidade relativa atinge os seus maiores valores nos meses de dezembro e janeiro e os mais baixos nos meses de julho e agosto (meses mais secos). Os ventos NO, N, SO e O, vindos do mar, tornam mais dolorosos os invernos, mas amenizam o clima durante as estações quentes. Os nevoeiros

são mais frequentes nos meses de agosto e de setembro, embora também possam aparecer nos meses que vão de outubro a março e em junho e julho [Prata, 2010].

Na Figura 3.2 está representado o gráfico da temperatura média e precipitação média mensal segundo as normais climatológicas para o período de 1981-2010, para a estação meteorológica de Setúbal (estação de referência mais próxima).

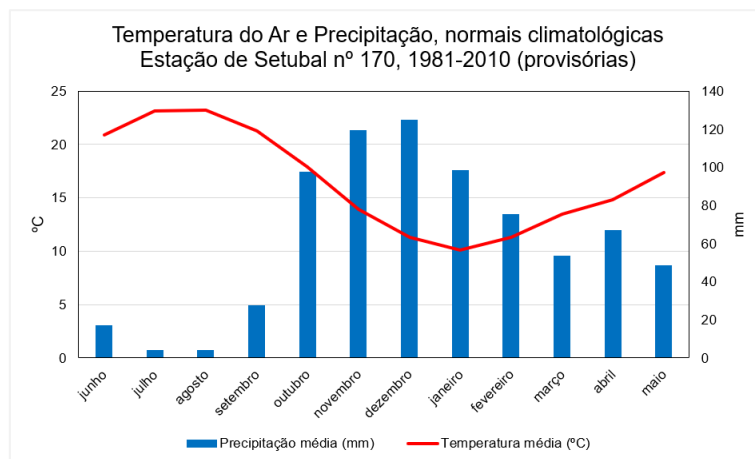


Figura 3.2 – Valores de precipitação média e temperatura média das normais climatológicas entre 1981 e 2010 para a estação de Setúbal [adaptado de IPMA, 2018].

3.2. DADOS UTILIZADOS

O processamento dos dados recolhidos exigiu a utilização de diversas aplicações informáticas, nomeadamente ESA SNAP®, ArcGIS 10.4®, ENVI 5.3®, MATLAB R2015a®, e Microsoft® e Excel 365 (Tabela 3.1).

Tabela 3.1 – Lista das aplicações informáticas utilizadas e respetiva função

Função	Aplicação Informática
<ul style="list-style-type: none"> Processamento das imagens Sentinel-2 	ESA SNAP
<ul style="list-style-type: none"> Cálculo das imagens de NDVI Tratamento de dados vetoriais (parcelas hortícolas) Cálculo de parâmetros estatísticos por parcela 	ArcGIS 10.4
<ul style="list-style-type: none"> Classificação e validação dos resultados 	ENVI 5.3
<ul style="list-style-type: none"> Desenvolvimento do algoritmo para análise píxel a píxel Criação de bandas artificiais 	MATLAB R2015a
<ul style="list-style-type: none"> Estudo do comportamento de NDVI por parcela hortícola Tratamento estatísticos dos resultados das classificações 	Excel 365

3.2.1. Imagens de Satélite

A análise da série temporal foi feita com base nas imagens selecionadas para o período em estudo, compreendido entre junho de 2017 e maio de 2018. Selecionaram-se 35 imagens dos satélites Sentinel-2A e Sentinel-2B (Tabela 3.2), disponibilizadas gratuitamente através da plataforma *online* Copernicus Open Access Hub da ESA (*European Space Agency*) e georreferenciadas no sistema UTM29N/WGS84. Apenas foi necessário garantir que estas abrangiam a Península de Setúbal, não estando afetadas por nuvens nem evidenciando efeitos atmosféricos (nevoeiro ou neblina) na área de estudo, possibilitando assim a análise multitemporal.

Tabela 3.2 – Data de aquisição das imagens Sentinel-2A e Sentinel-2B, sob correção atmosférica Nível-1C¹.

Ano	2017							2018				
Mês	jun	jul	ago	set	out	nov	dez	jan	fev	mar	abr	mai
	4	4	3	2	2	6	1	15	24	21	25	5
	14	9	8	12	7	11	16	30		26		
Dia		14	13	22	12	16	21					
		24	18	27	22	21						
		29	23		27							

As principais características das imagens da missão Sentinel-2 encontram-se listadas na Tabela 3.3. Da informação disponibilizada, foram utilizadas 2 das 13 bandas espectrais dos sensores MSI (Multispectral Instrument): B04 (Vermelho) e B08 (Infravermelho Próximo).

Tabela 3.3 – Principais características das imagens Sentinel-2.

Missão	Bandas Espectrais		Comprimento de Onda (µm)	Resolução Espacial (m)
Sentinel-2	B02	Blue	490	10
	B03	Green	560	
	B04	Red	665	
	B08	NIR	842	
	B05	Red Edge 1	705	20
	B06	Red Edge 2	740	
	B07	Red Edge 3	783	
	B8A	Red Edge 4	865	
	B011	SWIR 1	1610	
	B012	SWIR 2	2190	
	B01	Coastal Aerosol	443	60
	B09	Water Vapor	945	
	B010	Cirrus	1380	

A missão Sentinel-2 composta primeiramente pelo satélite 2A, lançado a 23 de junho de 2015, e, posterior e adicionalmente pelo satélite 2B, lançado a 7 de março de 2017, é uma missão europeia de observação da Terra para obtenção de imagens multiespectrais de alta resolução e com uma larga faixa de aquisição. Os dois satélites estão posicionados na mesma órbita (heliossíncrona polar), com um desfasamento de 180°, a 786 km de altitude e cobrindo uma faixa de 290 km de largura.

¹ As imagens estão ortorretificadas e os valores dos píxeis correspondem a valores de refletância relativos ao topo da atmosfera (TOA) [adaptado de ESA, 2018].

A resolução temporal é de 10 dias com um satélite, e de 5 dias com dois satélites operacionais. Cada imagem do satélite Sentinel-2 tem uma dimensão de 100 por 100 km, com 12 bits por píxel, e uma resolução espacial de 10, 20 ou 60 metros, consoante a banda. A missão Sentinel-2 (S2) garante a continuidade de aquisição de dados das missões Landsat e SPOT, destacando-se pela maior resolução espectral e espacial que apresenta [Drusch *et al.*, 2012].

3.2.2. Informação das Parcelas Hortícolas

Os dados relativos às parcelas hortícolas em estudo foram fornecidos pela PrimoHorta Sociedade de Produtores de Hortícolas, LDA. Esta organização de produtores tem como principais produtos hortícolas gerados a cenoura, batata, couve e cebola, associados a uma área de produção de cerca de 1000 hectares (10 km²), sendo responsáveis pelo abastecimento das grandes unidades comerciais do país, realizando periodicamente exportações para outros países da União Europeia.

A informação referente às 34 das 40 parcelas hortícolas fornecidas pela PrimoHorta foi obtida através de um ficheiro vetorial com o sistema de referência Hayford-Gauss *Datum* 73, posteriormente transformado para o sistema de referência UTM29N/WGS84, utilizado na georreferenciação das imagens de satélite (Figura 3.3). O ficheiro contém atributos que discriminam o tipo de cultura de cada parcela, bem como as respetivas datas de sementeira e colheita associadas (Anexo A). Estes dados foram utilizados na identificação das parcelas possibilitando assim a análise do comportamento fenológico das culturas.

As parcelas hortícolas apresentam uma área média de 7.49 ha, correspondendo a sua área na totalidade a 0.95% do *subset* em estudo. Das 40 parcelas cedidas, 6 foram excluídas por não apresentarem indícios de plantação, uma vez que o comportamento da curva do ciclo fenológico era duvidoso.

Tabela 3.4 – Percentagem de ocupação de cada cultura em estudo.

Parcelas	Face às culturas	Face à área de estudo
Batata	52.02 %	0.49 %
Cenoura	45.90 %	0.44 %
Couve	2.08 %	0.02 %
Área total (ha)	≈ 314	33 068

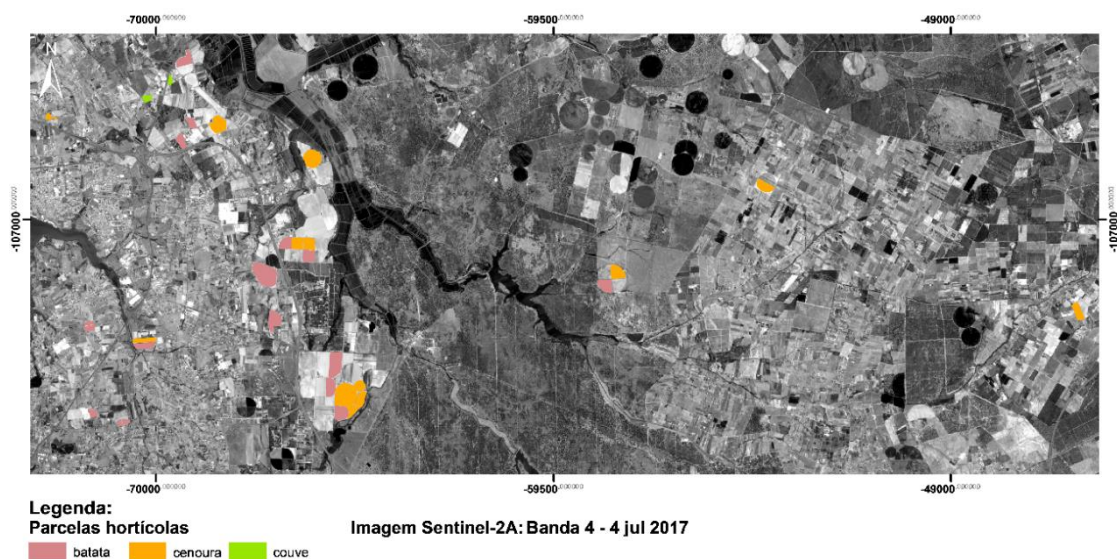


Figura 3.3 – Parcelas Hortícolas usadas no estudo da fenologia das culturas de batata, cenoura e couve.

3.2.3. Sistema de Identificação de Parcelas

O parcelário agrícola, também designado por Sistema de Identificação de Parcelas (SIP) é o registo na base de dados georreferenciada do Ministério da Agricultura e do Mar (MAM) das parcelas exploradas em agricultura e floresta. O MAM, através do Instituto de Financiamento da Agricultura e Pescas, I.P. (IFAP) e em conjunto com as Organizações de Agricultores, as Direções Regionais de Agricultura e Pescas e a Autoridade Florestal Nacional, procederam, em 2011, à revisão do parcelário de todas as explorações agrícolas. O SIP é um sistema de informação geográfica que tem como objetivo a identificação do limite das parcelas das explorações agrícolas, às quais é atribuído um único número, assim como a delimitação e classificação das ocupações de solo. Permite a referência geográfica das explorações agrícolas de modo unificado e coerente, simplificando os elementos gráficos que devem ser fornecidos pelos requerentes para fins de cálculo das Ajudas Comunitárias e para as Ações de Controlo. A obtenção do parcelário é indispensável para se apresentarem candidaturas aos apoios financeiros da União Europeia e do Estado Português [IFAP, 2018].

Para a obtenção dos dados de teste e de validação foi também utilizada a informação do SIP fornecida pelo IFAP, num total de 144 parcelas referentes ao período das imagens em estudo (Tabela 3.5). Esta informação foi disponibilizada em formato vetorial no sistema de coordenadas Hayford-Gauss *Datum* 73, também esta posteriormente transformada para o sistema de referência UTM29N/WGS84. A estes polígonos representativos das parcelas agrícolas foram ainda adicionados outros polígonos relativos a outros tipos de uso e ocupação do solo existentes na área de estudo, nomeadamente 10 polígonos de água e 30 polígonos correspondentes a diferentes coberturas artificializadas, tais como estrada, edifícios com telhado de telha, edifícios com telhado metálico, parques de estacionamento, entre outros (Figura 3.4).

Tabela 3.5 – Culturas controladas para dados de treino e validação.

Parcelas Controladas							Total: 184
Batata	Milho	Pastagem	Pousio	Prados	Tomate	Vinha	Total do IFAP
8	12	18	18	42	6	40	144

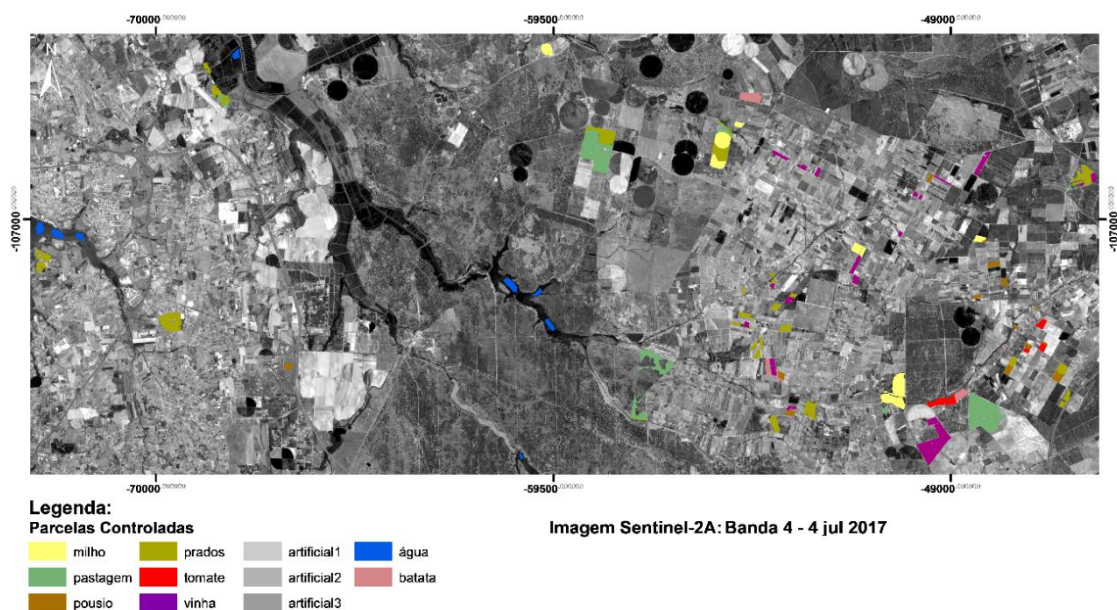


Figura 3.4 – Parcelas Hortícolas usadas para teste e validação da classificação.

3.3. MÉTODOS

A metodologia desenvolvida neste trabalho de projeto pretende, com recurso a tecnologias de DR, a identificação de culturas hortícolas ao longo do seu ciclo de crescimento. Como as hortícolas são culturas de irrigação com um ciclo fenológico conhecido, ou possível de ser caracterizado no tempo para cada espécie hortícola, propõe-se uma metodologia baseada na análise ao nível do píxel, da variabilidade temporal do NDVI. Nesta análise são incluídos os parâmetros que caracterizam cada cultura, tais como a duração do ciclo fenológico, a intensidade máxima, média, mínima e o desvio padrão do NDVI, bem como a amplitude da curva de NDVI.

A metodologia aplicada divide-se em 4 etapas principais: (1) tratamento das imagens Sentinel-2 e dos ficheiros vetoriais com as parcelas hortícolas, (2) estudo da fenologia das culturas, (3) desenvolvimento do algoritmo para a análise da variabilidade temporal do NDVI píxel a píxel, (4) e classificação e validação dos resultados obtidos (Figura 3.5).

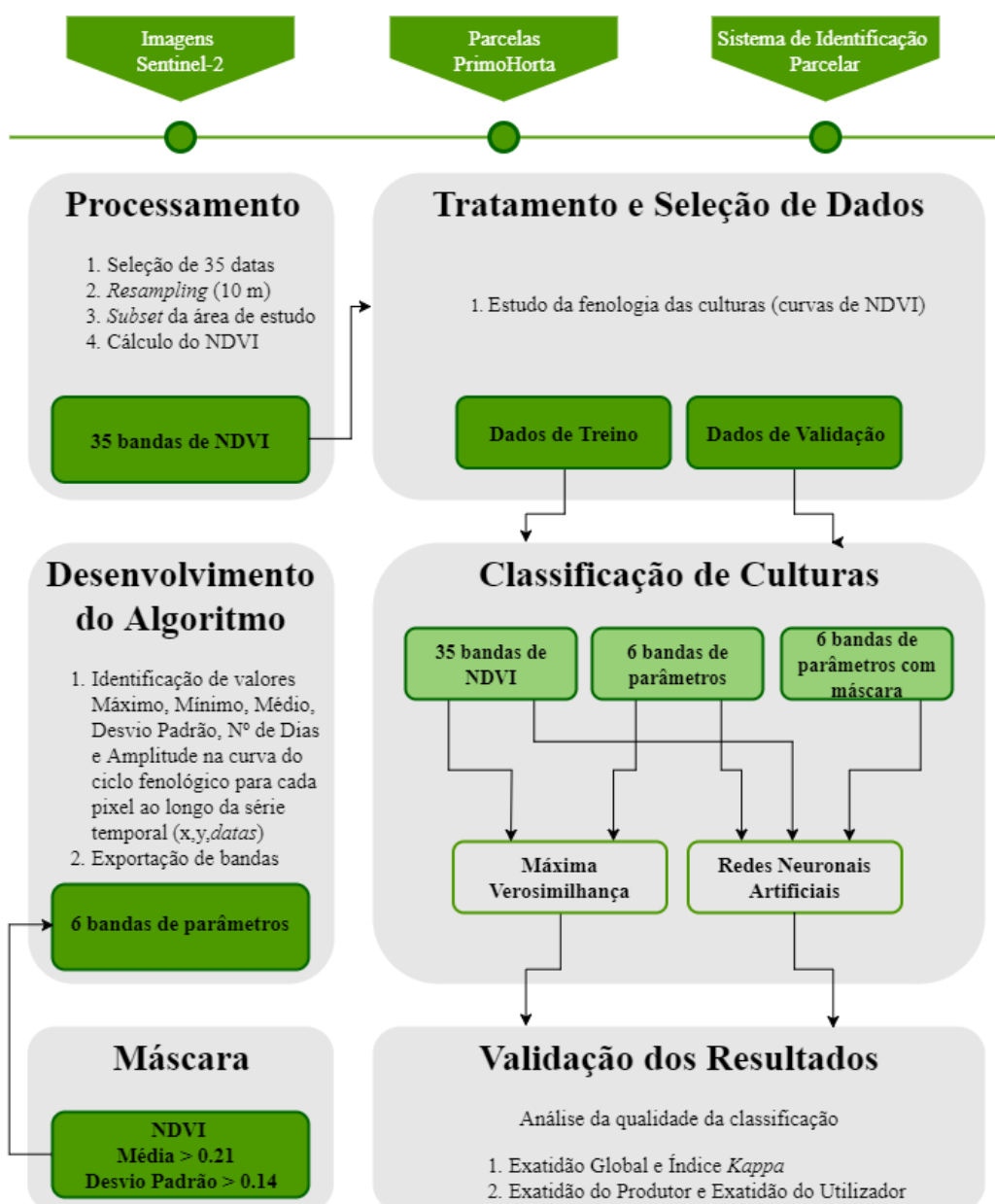


Figura 3.5 – Diagrama do trabalho de projeto.

3.3.1. Cálculo do Índice de Vegetação NDVI

Para cada uma das 35 datas de aquisição selecionadas foi feito o *download* de todas as imagens, tendo depois sido selecionadas as bandas B04 e B08, correspondentes ao comprimento de onda do vermelho (Red) e do infravermelho próximo (NIR) no espectro eletromagnético, respetivamente (Tabela 3.3). As mesmas foram agrupadas numa estrutura BSQ (*Band Sequential*) bem como “isoladas” à área de estudo através de um *subset* (1181 linhas por 2800 colunas), originando uma imagem com as combinações necessárias ao cálculo do NDVI.

A informação extraída das imagens aplicou-se especialmente à caracterização dos vários períodos fenológicos das culturas ao longo do seu ciclo de crescimento, com base no comportamento do NDVI ao longo da série temporal. Os índices de vegetação baseiam-se nas diferenças de refletividade que a vegetação de cor verde apresenta nas regiões do visível e do infravermelho. Têm dois objetivos principais, realçar aspetos que permitem discriminar entre diferentes tipos e estados de vegetação e minimizar radiancias do mesmo tipo de coberto, resultantes de diferenças de iluminação devidas à topografia [Fonseca e Fernandes, 2004]. A determinação dos valores de NDVI para cada uma das datas de aquisição das imagens foi realizada através da razão de diferença normalizada entre as bandas do infravermelho próximo e do vermelho [Rouse *et al.*, 1973], expressa na equação (1).

$$NDVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}} \quad (1)$$

Os valores NDVI variam entre -1 e 1, sendo que os valores negativos representam nuvens ou água, os valores ao redor de zero representam solo nu ou sem vegetação, e quanto mais alto for o valor de NDVI mais verde, densa e vigorosa está a vegetação.

Para o efeito, foi desenvolvido em linguagem Matlab® um programa que considera como dados de entrada as imagens S2 e um ficheiro vetorial com os limites das parcelas. Com base na junção desta informação, o programa calcula o valor de NDVI para cada parcela ao longo da série temporal de imagens devolvendo um ficheiro .txt com a informação do número da parcela, do número da imagem, do valor médio e do desvio padrão do NDVI para cada parcela em cada época de aquisição das imagens S2. Para além do ficheiro de saída o programa gera também duas imagens, uma com os valores de média temporal do NDVI e outra como os valores de desvio padrão temporal. Adicionalmente, foram criadas 35 imagens de NDVI correspondentes a cada uma das datas, através da ferramenta *ModelBuilder* do ArcGis®, tendo como designação (nome) a respetiva data.

3.3.2. Avaliação Fenológica: criação e estudo de Curvas de Crescimento por Parcela

No estudo da fenologia são identificadas as diversas fases do ciclo de crescimento das culturas, tais como a germinação, o florescimento e a maturação. Neste contexto, mediante a resposta espectral de cada cultura, estudou-se o modo como estas se comportavam ao longo da série temporal, avaliando a dinâmica de cada hortícola a partir dos máximos e mínimos dos valores de NDVI para cada curva representativa do ciclo de crescimento de cada cultura. Para o efeito, com base nos valores de NDVI calculados *a priori*, foram geradas as curvas de crescimento de cada parcela de batata, cenoura e couve em estudo, avaliando-se paralelamente a compatibilidade das datas de sementeira e colheita fornecidas pela PrimoHorto (Anexo B).

A partir da análise visual de cada uma das curvas e considerando-se a curva teórica da Figura 3.6, foram estabelecidos vários parâmetros para a caracterização de cada cultura:

- $NDVI_{Max}$: Valor máximo de NDVI para a data de maturação;
- $NDVI_{MinE}$: Valor mínimo de NDVI para a data de sementeira;
- $NDVI_{MinD}$: Valor mínimo de NDVI para a data de colheita;
- $NDVI_{Amplitude}$: Diferença entre a média dos valores mínimos e o valor máximo;
- $NDVI_{N^{\circ}Dias}$: Número de dias decorrentes entre $NDVI_{MinE}$ e $NDVI_{MinD}$.

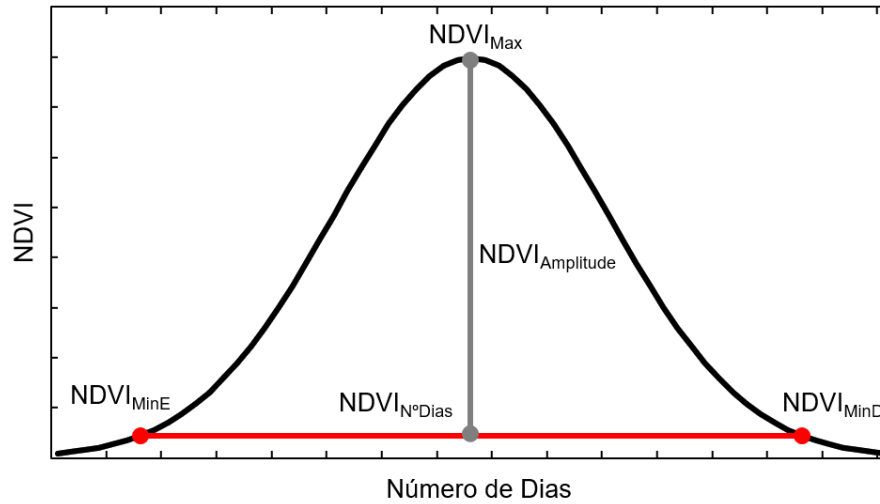


Figura 3.6 – Representação gráfica do ciclo de crescimento teórico das culturas hortícolas com os parâmetros considerados.

3.3.3. Desenvolvimento do Algoritmo

A classificação temporal pixel a pixel teve por base a criação de um algoritmo em linguagem *Matlab* que gerasse 6 bandas artificiais correspondentes ao valor do NDVI máximo, mínimo, médio, desvio padrão, amplitude da curva de crescimento e número de dias decorrentes entre a sementeira e a colheita, calculado para cada pixel ao longo da série temporal (Figura 3.7).

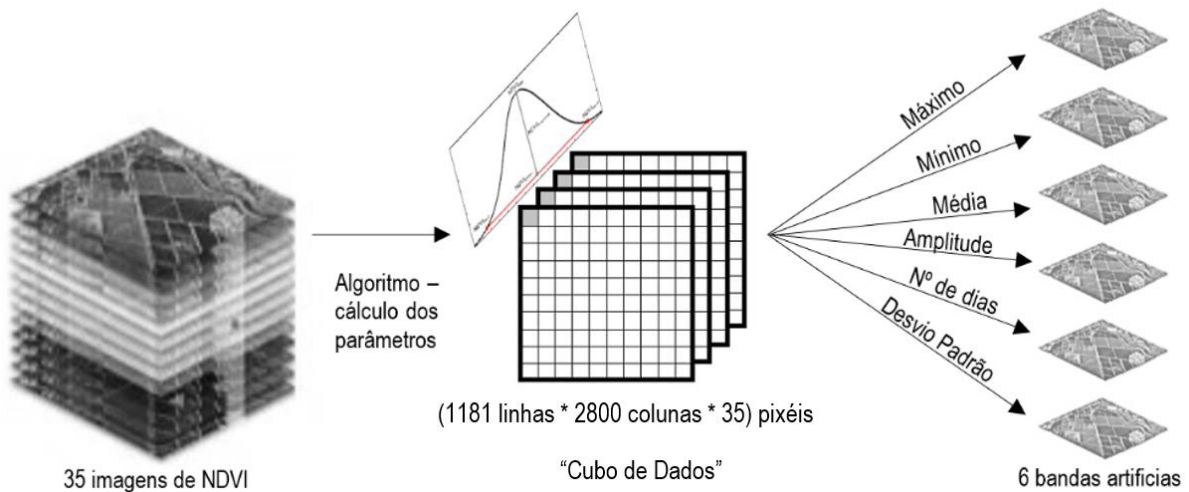


Figura 3.7 – Representação esquemática do algoritmo aplicado na criação das 6 bandas artificiais.

Foram testadas três abordagens distintas para o cálculo das bandas artificiais “Nº de dias” e “Amplitude” recorrendo à função *findpeaks* (Figura 3.8), enquanto que as restantes bandas artificiais foram calculadas com funções predefinidas do *Matlab* (*max*, *min*, *mean* e *std*).

Na Versão 1, numa primeira instância o algoritmo procura os máximos locais, seguindo-se o cálculo do valor máximo dos máximos, depois, sob inversão da função, prossegue com a pesquisa dos mínimos locais com valor de NDVI inferior a 0.3, por fim seleciona os mínimos locais mais próximos à esquerda e à direita do valor máximo encontrado. A amplitude da curva de crescimento da cultura é dada pela diferença entre o valor mínimo médio e o valor máximo. O número de dias do ciclo corresponde à diferença de posição (datas) entre os 2 valores mínimos de NDVI considerados. Adicionalmente, quando à direita de um pico máximo não existem mínimos (por não existirem mais imagens e por isso valores que completem o ciclo fenológico) o algoritmo considera o mínimo à esquerda multiplicando o número de dias desse mínimo até ao máximo por dois. Se não existirem valores mínimos à direita e à esquerda do máximo superiores a 0.25, é atribuído o valor de 0 para o número de dias e para a amplitude (geralmente para píxeis de água).

A Versão 2 difere da Versão 1 na medida em que não há inversão da função, pelo que os valores considerados são apenas os máximos locais. Considera-se o valor máximo dos máximos para o valor de maturação máximo do NDVI, os mínimos vão ser os máximos locais inferiores a 0.3 mais próximos do valor máximo. A amplitude e o número de dias são calculados de modo semelhante à Versão 1, no entanto esta abordagem conduz a um aumento significativo do número de dias, uma vez que os mínimos se encontram mais afastados do valor máximo.

Na Versão *Half* o algoritmo também procura os máximos locais, calculando depois o máximo dos máximos, no entanto, é considerado apenas um mínimo. Com base também na função inversa calculam-se os mínimos locais e aquele que estiver mais próximo do máximo será o escolhido. A amplitude é novamente calculada pela diferença de valores máximo e mínimo, porém, o número de dias é multiplicado por dois. Dada a maior proximidade do valor mínimo ao valor máximo, nesta versão o número de dias pode diminuir significativamente.

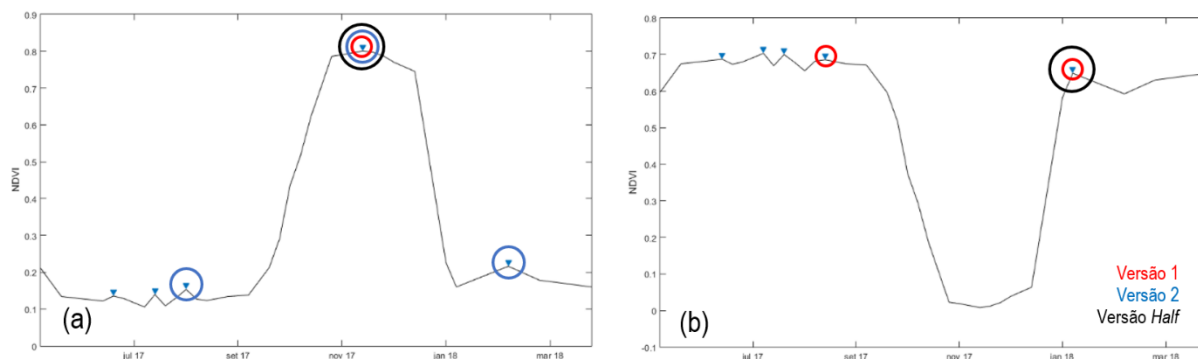


Figura 3.8 – Ilustração do comportamento da Versão 1, Versão 2 e Versão *Half* na escolha de máximos e mínimos para um mesmo píxel da série temporal, sendo (b) função inversa de (a).

3.3.4. Criação de uma Máscara

De forma a excluir zonas artificializadas e a melhorar a exatidão da classificação aplicou-se uma máscara para delimitar essencialmente vegetação, assumindo píxeis com desvio padrão de NDVI superior a 0.14 e média superior a 0.21 (Tabela 3.6). Esta delimitação teve por base os valores médios calculados por classe de parcela, tendo em conta a média do desvio padrão e a média das médias. Com esta delimitação consideraram essencialmente as zonas correspondentes a vegetação (culturas a classificar) com objetivo de melhorar a exatidão dos resultados.

Tabela 3.6 – Valores médios de NDVI por classe (média das médias e média do desvio padrão).

Classe	Média	Desvio Padrão
Água	-0.06	0.13
Artificial 3	0.02	0.04
Artificial 2	0.06	0.05
Artificial 1	0.11	0.04
Pousio	0.27	0.16
Vinha	0.28	0.09
Prados	0.34	0.14
Batata	0.34	0.21
Couve	0.35	0.23
Pastagem	0.38	0.11
Tomate	0.39	0.19
Cenoura	0.39	0.26
Milho	0.45	0.24
Valores considerados na máscara	0.21	0.14

3.3.5. Classificação e Validação

Os procedimentos de classificação de imagem identificam grupos homogêneos que representam a cobertura da superfície terrestre. Os algoritmos de classificação de imagem, com base na informação espectral representada pelos números digitais numa ou mais bandas espectrais, classificam cada píxel. O objetivo passa por atribuir classes ou temas a todos os píxeis de uma imagem, gerando um mapa temático da imagem original [CCRS, 2018].

A classificação supervisionada é um procedimento frequentemente utilizado para análises quantitativas de dados de DR. Esta utiliza algoritmos adequados para classificar os píxeis de uma imagem em classes, que pretendem representar diferentes tipos de uso e ocupação do solo. Existem diversos algoritmos disponíveis para este fim, contudo independentemente do método utilizado os principais passos do processo incluem: a seleção das classes a representar, a recolha dos dados de treino, a estimativa de parâmetros do classificador calculados com base nos dados de treino, a utilização do algoritmo para classificação dos píxeis da imagem nas classes desejadas, a produção de mapas temáticos ou tabelas de síntese dos resultados obtidos e a avaliação da exatidão do produto final utilizando dados de validação [Richards e Jia, 2006].

Para o estudo proposto neste trabalho de projeto foram testados dois classificadores, sendo eles o da Máxima Verosimilhança (MV) e o das Redes Neurais Artificiais (RNA). Para serem efetuados testes e retirar conclusões de quais os melhores resultados, foram consideradas diferentes combinações de bandas para a classificação das imagens: (a) 35 bandas de NDVI em estrutura BSQ; (b) 6 bandas artificiais em estrutura BSQ; (c) e 6 bandas artificiais com a máscara (ver seção 3.3.5) aplicada em estrutura BSQ.

3.3.5.1. Máxima Verosimilhança (MV)

O método da MV é um dos mais utilizados na classificação de imagens em DR. Tal como noutros métodos de classificação, no método da MV as classes espectrais conseguem ser descritas por uma distribuição de probabilidades no espaço multitemporal. Este método classifica os píxeis de acordo com a regra de Bayes que determina a probabilidade de um determinado píxel, através da sua resposta

espectral e com base na sua localização, pertencer a uma determinada classe, seguindo a distribuição normal gaussiana [Richards e Jia, 2006]. A probabilidade é dada pela equação (2):

$$p(w_i|x) = p(x|w_i) p(w_i) / p(x) \quad (2)$$

onde $p(w_i)$ é a probabilidade que cada classe w_i tem de ocorrer na imagem e x representa o valor do píxel em cada banda.

As vantagens de utilizar o classificador da MV resumem-se ao facto de este ter em conta a variância-covariância dentro das distribuições da classe e de este possuir uma performance mais elevada para dados normalmente distribuídos, que outros classificadores paramétricos. Contudo, para dados que não possuam uma distribuição normal os resultados podem não ser satisfatórios [Otukey e Blaschke, 2010].

3.3.5.2. Redes Neurais Artificiais (RNA)

As RNA são modelos estatísticos utilizados na área da aprendizagem automática (*machine learning*) e inspirados por redes neurais biológicas como as do sistema nervoso central. Uma rede neuronal é um algoritmo não paramétrico em camadas, composto por elementos de processamento, denominados de nós. O elemento principal de processamento tem em conta vários dados de entrada, dando origem a um único resultado. Geralmente, o número de dados de entrada é definido pela topologia da rede e pela dimensionalidade dos dados. Segundo Richards e Jia [2006], a operação associada a cada nó é descrita pela equação (3):

$$o = f(w^t x + \theta) \quad (3)$$

onde w é o vetor de pesos, x o vetor de dados de entrada (valores das bandas para um determinado píxel), θ o limite, e f a função de ativação.

Normalmente uma rede neuronal de multicamadas é projetada com uma camada de nós de entrada e uma camada de saída (Figura 3.9). Entre as camadas inicial e final, podem surgir uma ou mais camadas de nós ocultas ou de processamento. O número de camadas a utilizar deve ser seleccionado, sendo suficiente de um modo geral, uma rede de três camadas [Richards e Jia, 2006].

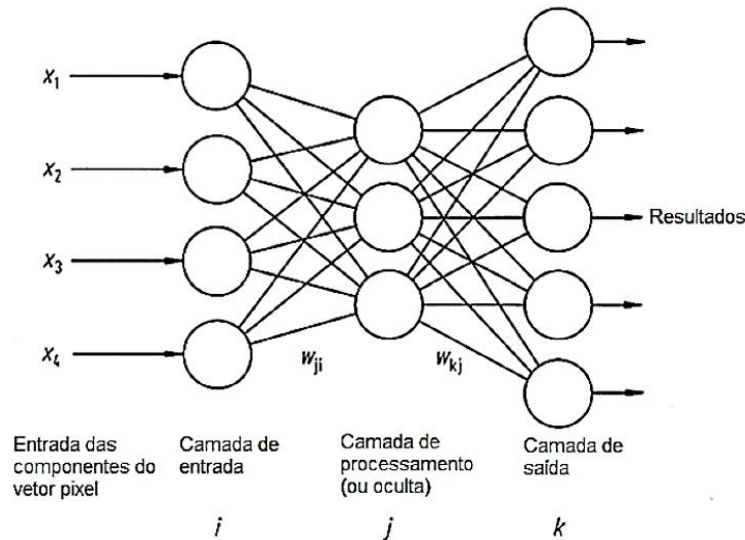


Figura 3.9 – Rede Neuronal de multicamadas. Adaptado de Richards e Jia [2006].

Utilizando os dados de treino, antes da classificação ser executada, a rede sofre um processo de aprendizagem supervisionada através da técnica de retropropagação (*backpropagation*). De uma forma geral, a aprendizagem consiste no ajustamento dos pesos ao nível do nó para minimizar a diferença entre

a ativação do nó de saída e o resultado. Outros dos parâmetros relevantes ao ajustamento da rede e passíveis de ser alterados são a taxa de treino, a taxa de momento de treino, o critério de paragem e o número de iterações. O erro é retropropagado através da rede e o ajustamento dos pesos é realizado através de um método recursivo.

3.3.5.3. Análise da Qualidade da Classificação

A fase final consistiu na avaliação da qualidade das classificações com base na tabela de contingências. Esta tabela, também designada por matriz de confusão, permite descrever as exatidões individuais de cada uma das classes identificadas no mapa. As colunas desta matriz representam as classes na amostra de referência, e as linhas representam as atribuições realizadas para cada classe da amostra, pelo procedimento utilizado. Com base nesta matriz pode calcular-se a Exatidão Global (EG), a Exatidão do Utilizador (EU) e a Exatidão do Produtor (EP). Um outro coeficiente que permite avaliar a exatidão da classificação é o índice *Kappa*, que avalia a concordância entre os dados de referência e os dados classificados. Na Tabela 3.7 encontram-se os valores estabelecidos por Landis e Koch [1977] para classificar, numa escala de desempenho muito fraco a desempenho muito bom, a qualidade da classificação obtida com base no índice *Kappa*.

Tabela 3.7 – Desempenho da classificação segundo os valores do índice *Kappa* [adaptado de Landis e Koch, 1977].

Índice de <i>Kappa</i> (%)	<0	0-20	21-40	41-60	61-80	81-100
Desempenho da Classificação	Muito Franco	Franco	Razoável	Moderado	Bom	Muito Bom

CAPÍTULO 4

RESULTADOS E DISCUSSÃO

4.1. AVALIAÇÃO DA FENOLOGIA DAS CULTURAS

4.1.1. Análise da Evolução Temporal do NDVI para culturas hortícolas

A utilização de técnicas de DR aplicadas à inventariação hortícola fornece informação rápida e espacializada sobre a dimensão, o vigor e outros parâmetros biofísicos das culturas. Deste modo, foi possível efetuar o estudo fenológico para as 3 culturas em causa, batata, cenoura e couve.

O início do ciclo fenológico das culturas de irrigação é caracterizado por um baixo valor de NDVI com um aumento gradual ao longo do período de crescimento da planta (e aumento da área foliar) até atingir um máximo, geralmente igual ou superior a 0.7, designando-se por fase de maturação. De seguida segue-se a fase de diminuição da atividade fotossintética que corresponde a uma diminuição rápida do valor de NDVI até valores próximos de 0.3, associados à fase de senescência, ou seja, ao envelhecimento da planta.

Nas Figuras Figura 4.1(a), 4.2(a) e 4.3(a) apresentam-se os gráficos de evolução temporal do NDVI da amostra total de parcelas para as 3 culturas em causa. Verifica-se que o período de estudo abrange na totalidade todas as fases fenológicas de crescimento, desde a plantação até ao final da maturação das culturas. O crescimento rápido das mesmas está associado ao aumento dos valores de NDVI e a maturação à estabilização destes valores. Adicionalmente, todos os casos em que as culturas das diversas parcelas se encontram em fase de senescência/colheita, se caracterizam pela diminuição do NDVI, sendo inclusive possível verificar o início de novas plantações em algumas parcelas.

Relativamente à cultura de batata (Figura 4.1), verifica-se que a maioria das parcelas apresenta um comportamento semelhante ao longo do período em estudo. Da análise do gráfico, verifica-se que as fases de crescimento rápido terão ocorrido entre finais de agosto e meados de setembro e que o período de colheita se tenha dado essencialmente em dezembro, sendo que em algumas das parcelas as culturas foram colhidas até fevereiro. Estas oscilações devem-se ao facto de os produtores serem diferentes e por isso não semearem/colherem exatamente nas mesmas épocas. Por norma as culturas de batata apresentam um ciclo vegetativo de 90 a 110 dias, sendo bem mais longo em países de clima temperado, podendo inclusive chegar até 150 dias [Rodrigues *et al.*, 2009]. Das 16 parcelas de batata em estudo, destacam-se 3 valores importantes, 0.70 como média dos valores máximos de NDVI, 0.53 como amplitude média da curva e um valor médio de 118 dias desde a sementeira até à colheita (Tabela 4.1, para mais detalhes consultar o Anexo C).

Tabela 4.1 – Valores Médios calculados para o ciclo fenológico das 16 parcelas de Batata em *Excel*.

Curva do Ciclo Fenológico	Max	Amp _{Esq}	Amp _{Dir}	Amp _{Média}	Nº dias
Valor Médio	0.70	0.56	0.50	0.53	118

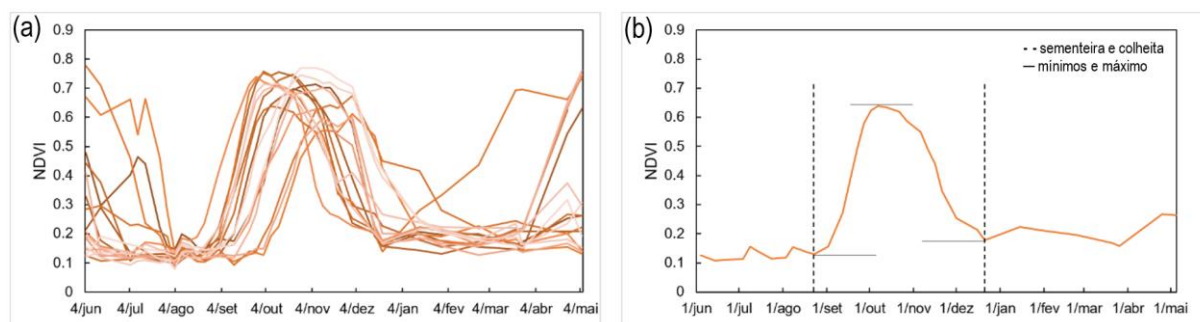


Figura 4.1 – (a) Evolução temporal do NDVI de todas as parcelas de batata estudadas. (b) Evolução temporal do NDVI para uma parcela genérica de batata (05-Quinta da Vinha).

Relativamente às culturas de cenoura (Figura 4.2), torna-se mais difícil de verificar se estas apresentam um comportamento semelhante ao longo do período em estudo, uma vez que as datas de sementeira diferem consideravelmente. Estas diferenças devem-se mais uma vez ao facto dos produtores serem distintos e por isso não semearem/colherem nas mesmas datas. Da análise do gráfico, supõe-se que as fases de crescimento rápido tenham tido início entre meados de julho e início de setembro e que o período de colheita tenha decorrido essencialmente entre dezembro e janeiro. Dependendo do cultivo, e das condições do clima, teoricamente a colheita da cenoura pode ser feita de 90 a 130 dias após a sementeira. Das 16 parcelas de cenoura em estudo, destacam-se 3 valores importantes, 0.74 como média dos valores máximos de NDVI, 0.57 como amplitude média da curva e um valor médio de 152 dias entre o período de sementeira e colheita (Tabela 4.2, para mais detalhes consultar o Anexo C).

Tabela 4.2 – Valores Médios calculados para o ciclo fenológico das 16 parcelas de Cenoura em *Excel*.

Curva do Ciclo Fenológico	Max	Amp _{Esq}	Amp _{Dir}	Amp _{Média}	Nº dias
Valor Médio	0.74	0.60	0.55	0.57	152

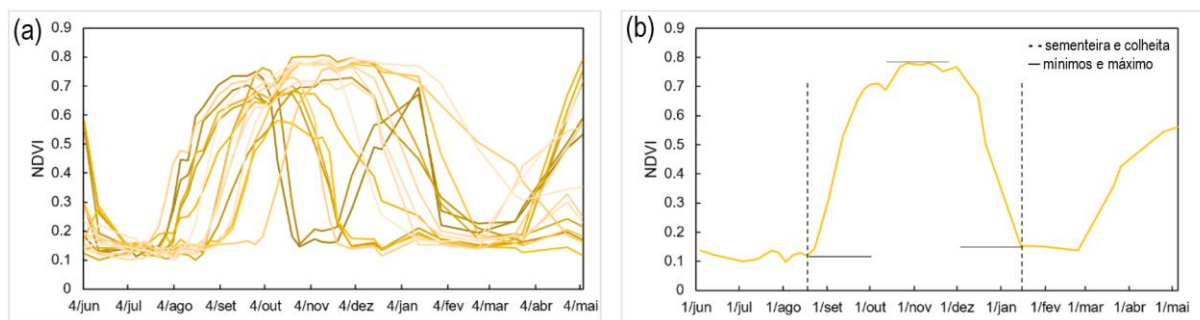


Figura 4.2 – (a) Evolução temporal do NDVI de todas as parcelas de cenoura estudadas. (b) Evolução temporal do NDVI para uma parcela genérica de cenoura (09B-Valmet).

A fraca representatividade das culturas de couve (Figura 4.3) não permite tirar grandes ilações, nem assumir que as mesmas apresentam um comportamento semelhante ao longo do período em estudo uma vez que só existiam 2 parcelas. Para esta cultura verifica-se numa das parcelas um período de senescência mais precoce face à outra, isto porque geralmente para culturas de couve a colheita pode variar entre 85 e 130 dias após a sementeira. Da análise da Tabela 4.3 destacam-se 3 valores importantes sendo eles, 0.72 como média dos valores máximos de NDVI, 0.54 como amplitude média da curva e um valor médio de 130 dias entre o período de sementeira e colheita (para mais detalhes consultar o Anexo C).

Tabela 4.3 – Valores Médios calculados para o ciclo fenológico das 2 parcelas de Couve em Excel.

Curva do Ciclo Fenológico	Max	Amp _{Esq}	Amp _{Dir}	Amp _{Média}	Nº dias
Valor Médio	0.72	0.56	0.52	0.54	130

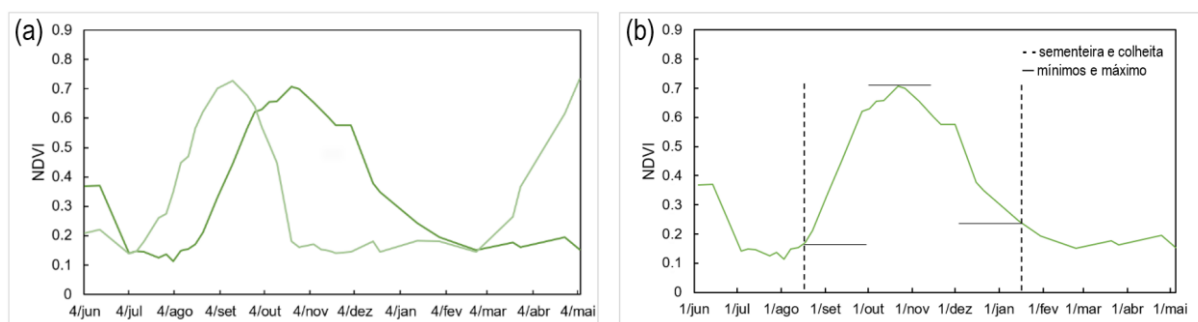


Figura 4.3 – (a) Evolução temporal do NDVI de todas as parcelas de couve estudadas. (b) Evolução temporal do NDVI para uma parcela genérica de couve (16-Rotunda).

Na Figura 4.4 encontram-se representadas as linhas de tendência média do NDVI, calculadas através da média de valores para cada data, para cada uma das culturas. Verifica-se que as 3 culturas apresentam valores máximos semelhantes, variando essencialmente no número de dias decorrente entre o mínimo de sementeira e o mínimo de colheita. Adicionalmente, é possível constatar que a cultura de cenoura tem um ciclo de duração superior ao da batata (devido ao desfasamento das culturas), bem como um período de senescência mais gradual em termos de descida do valor de NDVI. A avaliação do comportamento do ciclo fenológico de cada cultura permitiu a caracterização da mesma através do estudo de parâmetros como o número de dias e a amplitude associados a cada perfil temporal de NDVI. Esta interpretação visual e empírica permitiu ter uma noção da variação comportamental de cada curva, sendo posteriormente mais fácil ajustar o algoritmo na tomada de decisão face ao número de dias a considerar (mediante a escolha de valores máximos e mínimos).

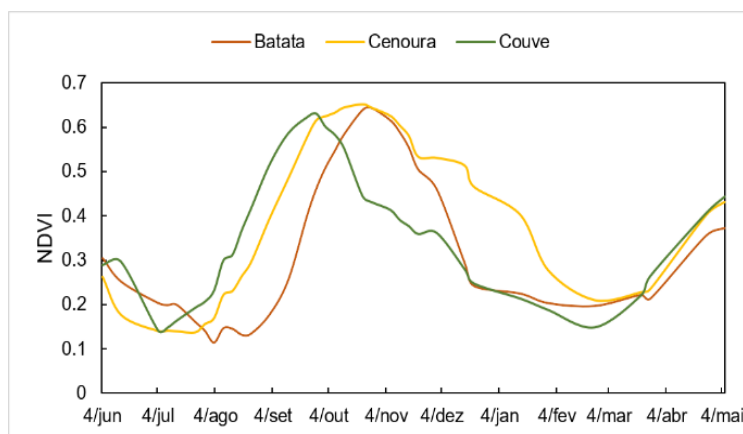


Figura 4.4 – Tendência média do NDVI para cada tipo de cultura.

4.1.2. Evolução Temporal do NDVI para culturas do IFAP

Adicionalmente foram geradas todas as curvas do ciclo fenológico das culturas fornecidas pelo IFAP para perceber se estas tinham um comportamento semelhante ao longo da série temporal, bem como para perceber se se assemelhavam às culturas hortícolas em estudo. Na Figura 4.5 é apresentada a evolução temporal do NDVI para as culturas de tomate, milho, vinha, prados, pastagem e pousio. Verifica-se que nas pastagens, nos prados e nos pousios o NDVI tem uma variabilidade sazonal com

valores de NDVI máximos essencialmente no inverno enquanto que as culturas de tomate, milho e vinha podem oscilar na época, mas têm um ciclo crescimento e decaimento de NDVI bem delimitado no tempo. Uma vez que a amostra era consistente e que as curvas das seis culturas não se equiparavam às de batata, cenoura e couve, pôde-se prosseguir com o algoritmo para classificação.

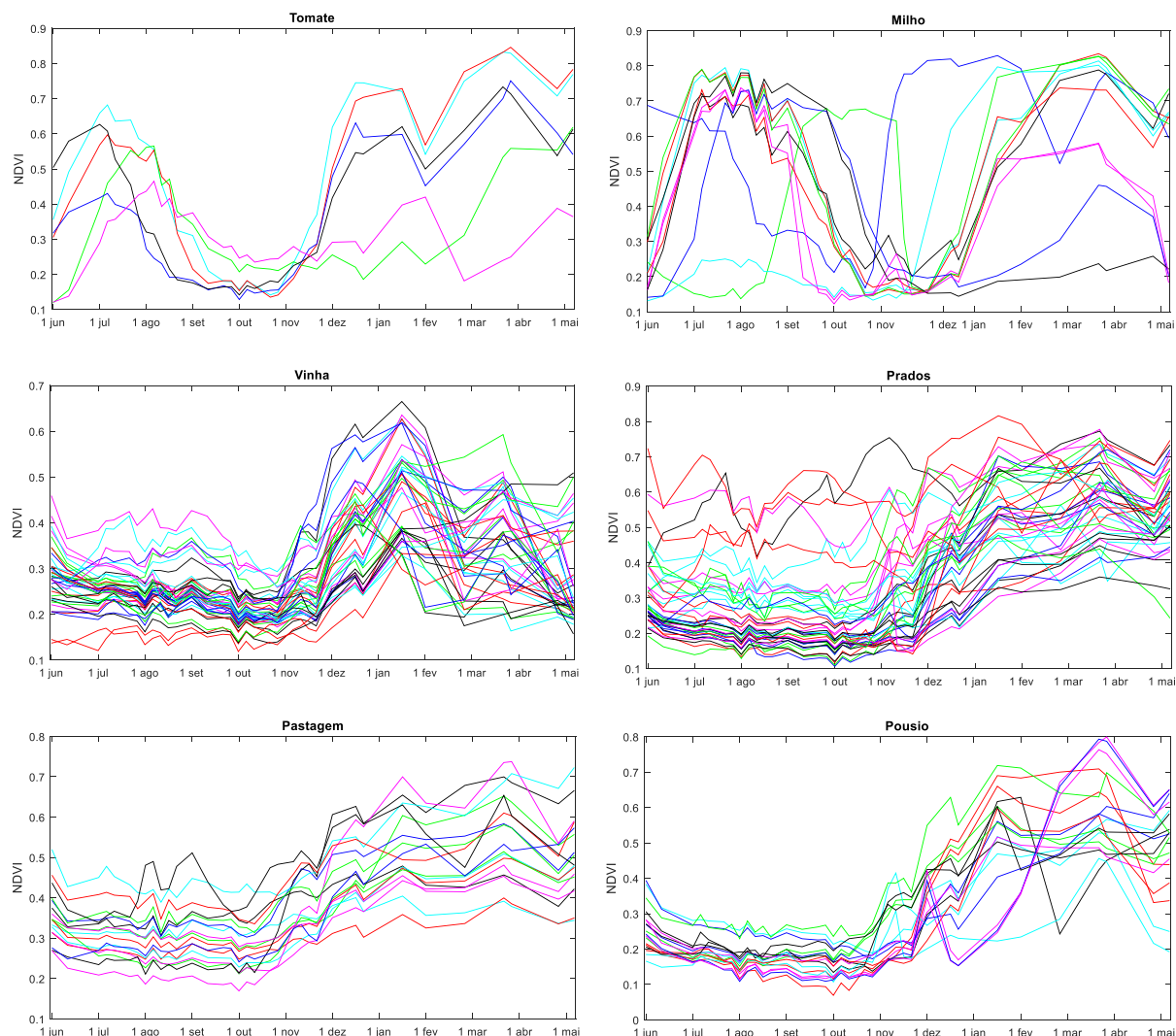


Figura 4.5 – Ciclo Fenológico das culturas de validação fornecidas pelo IFAP.

4.2. ALGORITMO PARA DETERMINAÇÃO DOS PARÂMETROS CARACTERÍSTICOS DA CURVA DE NDVI

4.2.1. Bandas Artificiais geradas

Considerando que um dos objetivos deste estudo consistia na caracterização de culturas hortícolas através dos parâmetros característicos de cada curva de NDVI, calculados pixel a pixel com base em todas as imagens da série temporal (cubo de dados), de seguida apresentam-se os resultados obtidos para cada um dos parâmetros.

A cada parâmetro corresponde uma imagem de números reais, resultando assim num conjunto de 6 imagens que foram designadas por bandas artificiais. Das três abordagens distintas testadas, o algoritmo

escolhido (Versão 1) foi aquele que apresentou menores diferenças comparativamente aos valores calculados com base na interpretação visual dos gráficos em *Excel* para o número de dias (Anexo D).

As bandas com o valor máximo, mínimo, médio e o desvio padrão foram calculadas através de funções do *Matlab* onde a margem de erro é nula, no entanto, as correspondentes ao número de dias e à amplitude apresentam valores que variam ligeiramente quando comparados com os calculados *a priori* em *Excel*. A Tabela 4.4 representa os parâmetros médios por hortícola para as 6 bandas artificiais desenvolvidas. Estes valores têm por base os parâmetros médios calculados para cada parcela pela Versão 1 do algoritmo (Anexo E).

Tendencialmente, os valores calculados através do algoritmo aumentam o número de dias uma vez que a sensibilidade do mesmo diminui para curvas onde os mínimos não estão tão bem definidos. Verifica-se ainda que os píxeis correspondentes a pastagem e prado, por terem um ciclo relativamente longo, uma amplitude pequena e ligeiras oscilações no decorrer do mesmo, apresentam valores de dias exageradamente elevados, sendo considerados valores erráticos (*outliers*) que posteriormente geram ruído aquando a classificação. Os valores de amplitude média da curva do ciclo fenológico são complexos de avaliar, uma vez que estão intrinsecamente relacionados com a diferença da posição dos mínimos locais calculados, no entanto, comparando as Tabelas Tabela 4.1, 4.2 e Tabela 4.3 com a Tabela 4.4 podemos observar que se mantém a mesma ordem de grandeza, tendo a cenoura uma amplitude de curva maior, seguida da couve e por fim da batata.

Tabela 4.4 – Valores médios calculados por classe de parcela através da versão 1 do algoritmo em MATLAB®.

Parcelas	Média	Desvio Padrão	NDVI máximo	NDVI mínimo	Amplitude da curva	Número de dias
Batata	0.33	0.22	0.73	0.11	0.55	158
Cenoura	0.39	0.26	0.78	0.11	0.60	170
Couve	0.35	0.23	0.77	0.11	0.56	117

As figuras que se seguem (Figura 4.6 a 4.11) correspondem às bandas artificiais correspondentes a cada parâmetro calculado pela Versão 1 do algoritmo.

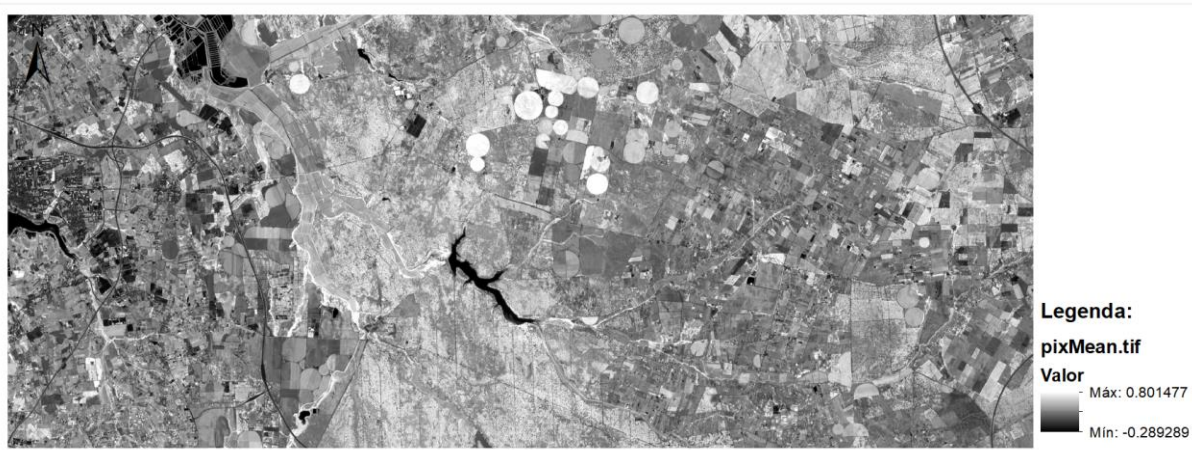


Figura 4.6 – Banda artificial representativa do valor médio e NDVI para cada píxel ao longo da série temporal, calculado através da função *mean* do MATLAB®, embutida na Versão 1 do Algoritmo.



Figura 4.7 – Banda artificial representativa do desvio padrão de NDVI para cada píxel ao longo da série temporal, calculado através da função *std* do MATLAB®, embutida na Versão 1 do Algoritmo.



Figura 4.8 – Banda artificial representativa do valor mínimo de NDVI para cada píxel ao longo da série temporal, calculado através da função *min* do MATLAB®, embutida na Versão 1 do Algoritmo.



Figura 4.9 – Banda artificial representativa do valor máximo de NDVI para cada píxel ao longo da série temporal, calculado através da função *max* do MATLAB®, embutida na Versão 1 do Algoritmo.



Figura 4.10 – Banda artificial representativa do número de dias para cada píxel ao longo da série temporal, calculados através da Versão 1 do algoritmo em MATLAB®.



Figura 4.11 – Banda artificial representativa da amplitude da curva para cada píxel ao longo da série temporal, calculados através da Versão 1 do algoritmo em MATLAB®.

4.2.2. Máscara com Zonas de Ocupação Hortícola

De modo a reduzir a confusão na classificação entre os vários tipos de classes, aplicou-se uma máscara às imagens sob garantia que se delimitavam as culturas em estudo, excluindo essencialmente áreas artificializadas. Com base nos valores da Tabela 3.7 foi criado um gráfico de média e desvio padrão para analisar a separabilidade das diferentes classes de uso e ocupação do solo (Figura 4.12). Assumindo-se como limite um desvio padrão inferior a 0.14 e uma média inferior a 0.21 criou-se uma máscara binária que permitiu uma redução de cerca de 71% no número de píxeis a classificar, eliminando assim, entre outras, áreas que não caracterizavam vegetação. Os píxeis foram identificados com o valor NaN (*Not a Number*) o que reduziu significativamente a quantidade dos mesmos a processar (Anexo F).

Verifica-se que os píxeis correspondentes a terreno artificializado, água e até vinha se diferenciam facilmente dos outros pelo desvio padrão, uma vez que os respetivos valores de NDVI oscilam menos ao longo da série temporal, como seria aliás de esperar.

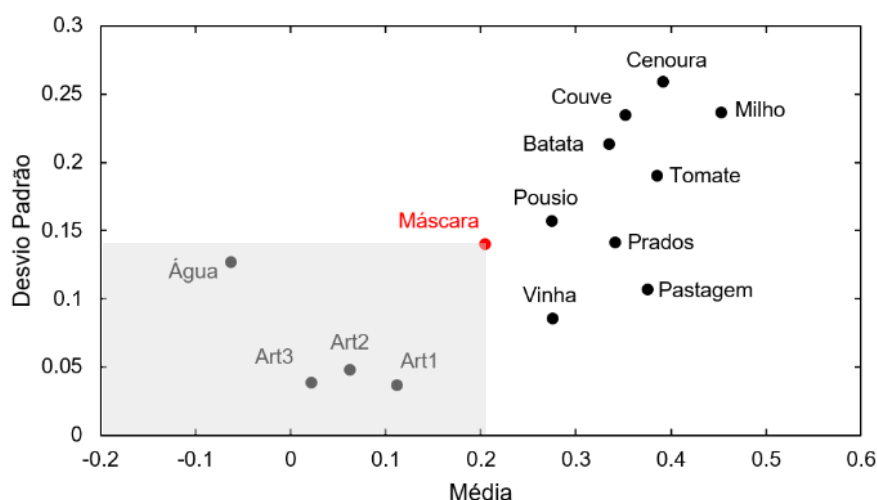


Figura 4.12 – Valores médios da Média e Desvio Padrão de NDVI, por classes, ao longo da série temporal.

4.2.2.1. Composição Colorida

Para além das imagens representadas em diferentes tonalidades de cinzento, correspondentes às refletâncias captadas para cada intervalo radiométrico, existem ainda imagens coloridas, resultantes da combinação das 3 componentes primárias de luz - as bandas vermelha (*Red*), verde (*Green*) e azul (*Blue*), vulgarmente designada por RGB. Estas imagens resultam do processo designado por composição colorida, em que as intensidades registadas para cada banda são representadas pela respetiva cor. Esta combinação de bandas consiste numa seleção cuidadosa que contém informações espectrais desejadas e que permita evidenciar cores e características para a percepção do olho humano [Liu e Mason, 2009].

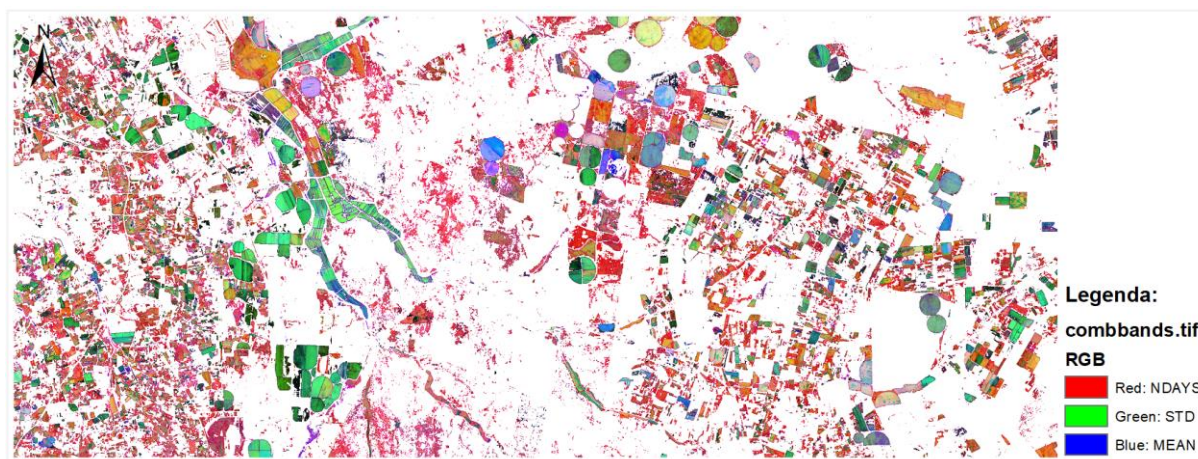


Figura 4.13 – Composição colorida das bandas número de dias (R), desvio padrão (G) e média (B) após a aplicação da máscara binária.

No contexto do projeto foi adicionalmente realizada uma composição colorida que teve como fonte as bandas do número de dias, do desvio padrão e da média calculadas através do algoritmo, com objetivo de perceber se as parcelas variavam entre culturas diferentes mantendo um padrão para culturas iguais. A geração desta imagem em falsa-cor permitiu salientar aspetos particulares das superfícies em estudo, nomeadamente no que concerne à tonalidade face ao comportamento das diferentes parcelas hortícolas, como se pode ver na Figura 4.14. As parcelas de batata geralmente aparentam um tom mais escuro e azulado em comparação com as de cenoura, já mais claras e verdes. Face ao número reduzido de parcelas de couve, é difícil tirar conclusões relativamente ao comportamento das mesmas, no entanto, a coloração

parece variar pouco comparando com as anteriores. Na imagem é ainda possível distinguir parcelas de tomate (tom rosa) e de milho (azul turquesa), entre outras.

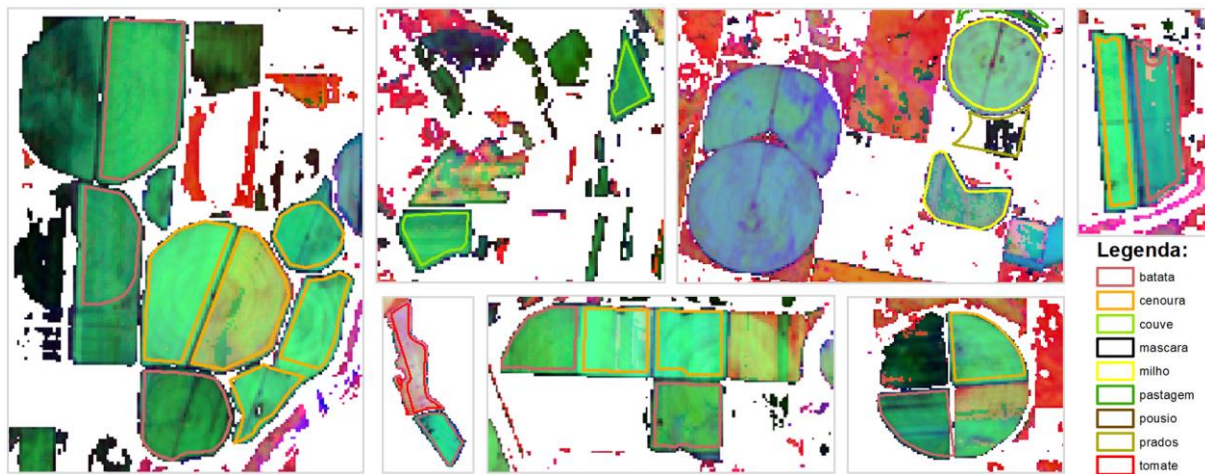


Figura 4.14 – Detalhe da composição colorida com algumas das parcelas em estudo.

4.3. CLASSIFICAÇÃO SEGUNDO OS MÉTODOS DE MÁXIMA VEROSIMILHANÇA (MV) E REDES NEURONAIS ARTIFICIAIS (RNA)

Nesta seção são apresentados e analisados os resultados obtidos no âmbito da classificação e validação das culturas em estudo. Primeiramente são apresentados e analisados os resultados dos métodos da MV e das RNA para a classificação multitemporal das culturas, com base nas 35 bandas de NDVI (como se de bandas espectrais se tratassem) e com base nas 6 bandas de parâmetros artificiais. Por fim são apresentados e analisados os resultados obtidos pelo método das RNA no estudo da parametrização do modelo, aquando a classificação das 6 bandas de parâmetros artificiais, com máscara.

4.3.1. Comparação entre MV e RNA

A primeira análise consistiu na comparação dos resultados da classificação obtidos através do classificador tradicional Máxima Verosimilhança com as Redes Neurais Artificiais, tanto usando as 35 bandas de NDVI como as 6 bandas artificiais representativas dos parâmetros.

A Tabela 4.5 contém o número de polígonos que se consideraram como dados de treino e dados de validação para cada uma das classes definidas. No presente estudo não foi possível recolher pelo menos 50 amostras por cada classe, como sugerido por Congalton e Green [2009], devido ao número diminuto de parcelas disponibilizadas pelo IFAP. As parcelas utilizadas como dados de treino e dados de validação podem ser consultadas no Anexo G.

Tabela 4.5 - Número de Parcelas de Treino (T) e de Validação (V) consideradas na comparação entre MV e RNA.

	Água	Art	Batata	Cenoura	Couve	Milho	Pastagem	Pousio	Prados	Tomate	Vinha
T	5	15	12	8	1	6	9	9	21	3	20
V	5	13	12	8	1	6	9	9	19	2	20

Relativamente ao algoritmo MV utilizaram-se os parâmetros por defeito sugeridos, sendo eles um valor único que define a área limite de probabilidade de todas as classes e um fator de escala igual a 1. Quanto ao algoritmo RNA assumiu-se por defeito que as 1000 iterações permitiam estabelecer um bom compromisso entre o tempo de processamento e a exatidão. Os restantes parâmetros foram também estabelecidos através dos valores sugeridos por definição, utilizando-se uma rede de três camadas, 1 camada oculta, uma taxa de treino de 0.2, uma taxa de momento de treino de 0.9 e um critério de paragem (erro médio quadrático) de 0.1.

Na Tabela 4.6 apresentam-se as exatidões globais e o índice *Kappa* obtidos para cada uma das combinações testadas com os métodos MV e RNA.

Tabela 4.6 – Resultados da exatidão global e do índice *Kappa* na comparação entre MV e RNA.

Método de Classificação	Exatidão Global (%)		Índice <i>Kappa</i> (%)	
	35 bandas de NDVI	6 bandas de parâmetros	35 bandas de NDVI	6 bandas de parâmetros
MV	67.29	51.37	61.62	44.88
RNA	21.90	39.62	16.97	29.73

Através da análise da tabela anterior, constata-se que o valor obtido para o índice *Kappa* aquando a classificação pelo método da MV para as 35 bandas é superior a 61%, o que corresponde a um bom desempenho segundo o critério da Tabela 3.7, todavia, analisando os restantes valores, estes variam entre moderado, razoável e inclusive franco. A exatidão global, que corresponde à percentagem de elementos corretamente classificados, variou consideravelmente entre os classificadores usados.

Genericamente, podemos dizer que o método MV teve melhor prestação para as 35 bandas de NDVI e que o método RNA apresentou melhores resultados para as 6 bandas artificiais de parâmetros.

A Tabela 4.7 e a Tabela 4.8 sintetizam as estatísticas alusivas à classificação das parcelas hortícolas em estudo para as 35 bandas de NDVI e 6 bandas de parâmetros artificiais, respetivamente. A avaliação da qualidade da classificação baseia-se numa tabela de contingências (matriz de confusão) que compara uma amostra dos píxeis classificados com a verdade no terreno, da qual se tiraram ainda indicadores como a exatidão do produtor, exatidão do utilizador, erro de omissão e erro de comissão.

A exatidão do produtor é obtida contabilizando os elementos da imagem que foram corretamente atribuídos a uma dada classe face ao número total de elementos que na amostra de referência foram atribuídos a essa classe (erros de omissão), já a exatidão do utilizador é obtida contabilizando os elementos da imagem que foram corretamente atribuídos a uma dada classe face ao número total de elementos que no mapa foram atribuídos a essa classe (erros de comissão). Logo, existem dois tipos de erro, o erro de omissão onde píxeis de uma determinada classe não são atribuídos a essa mesma classe, e os erros de comissão quando um píxel associado a uma classe na realidade pertence a outra.

Os resultados da classificação para cada cultura individual exibiram diferentes tipos de comportamento em termos de exatidão, dependendo do conjunto de bandas utilizado e do número de parcelas, bem como do classificador utilizado.

Tabela 4.7 – Resultados estatísticos para couve, cenoura e batata aquando o uso das 35 bandas de NDVI.

Matriz de Confusão (%)								
	MV			RNA				
	Couve	Cenoura	Batata	Couve	Cenoura	Batata		
Couve	0	0	0	0	0	0		
Cenoura	100	98.47	19.92	100	86.38	15.31		
Batata	0	0.99	78.62	0	0	0		

Exatidão da Classificação (%)								
	Exatidão Produtor		Erro Omissão		Exatidão Utilizador		Erro Comissão	
	MV	RNA	MV	RNA	MV	RNA	MV	RNA
Couve	0	0	100	100	0	0	0	0
Cenoura	98.47	86.38	1.53	13.62	66.31	47.82	33.69	52.18
Batata	78.62	0	21.38	100	91.64	0	8.36	0

Analisando a Tabela 4.7, alusiva à classificação aplicada às 35 bandas de NDVI, observamos que o erro de omissão das culturas de couve é de 100% para os dois métodos de classificação, o que seria de esperar face ao reduzido número de dados de treino disponibilizados. No que concerne à cultura de cenoura, estas apresentam para ambos os métodos uma exatidão do produtor bastante boa, de 98.47% através da MV e de 86.38% pelas RNA. Em relação à cultura de batata, a exatidão do utilizador é satisfatória para o algoritmo MV (78%), no entanto, esta cultura é omitida quando se aplica o algoritmo RNA.

Relativamente aos resultados das classificações aplicadas nas 6 bandas artificiais (Tabela 4.8), genericamente podemos assumir que estes ficaram aquém do esperado. Desta vez, apenas o método das RNA não classificou as culturas de couve, estando todas as outras classificadas, no entanto, os valores de exatidão do produtor rondam os 50%, existindo grande confusão entre as culturas.

Tabela 4.8 – Resultados estatísticos para couve, cenoura e batata aquando o uso das 6 bandas artificiais de parâmetros.

Matriz de Confusão (%)						
	MV			RNA		
	Couve	Cenoura	Batata	Couve	Cenoura	Batata
Couve	54.17	2.40	17.34	0	0	0
Cenoura	1.19	56.55	14.32	0	48.62	3.12
Batata	44.64	13.72	54.82	100	27.49	85.96

Exatidão da Classificação (%)								
	Exatidão Produtor		Erro Omissão		Exatidão Utilizador		Erro Comissão	
	MV	RNA	MV	RNA	MV	RNA	MV	RNA
Couve	54.17	0	45.83	100	5.63	0	94.37	0
Cenoura	56.55	48.62	43.45	51.38	67.42	82.38	32.58	17.62
Batata	54.82	85.96	45.18	14.04	77.48	61.95	22.52	38.05

Tendo em conta as semelhanças que o ciclo vegetativo da cenoura e da batata apresentam, bem como a proximidades dos valores médios e do desvio padrão na Figura 4.12, é possível constatar que a separação perfeita destas duas culturas em duas classes distintas é afetada por evidenciarem comportamentos semelhantes. Dado isto, decidiu-se agrupar os píxeis classificados como hortícolas (batata e cenoura) numa única classe, tendo neste caso sido obtidos valores de exatidão do produtor de 98.89% para o algoritmo MV e de 84.23% para o algoritmo RNA (Tabela 4.9).

Tabela 4.9 – Exatidões do Produtor e do Utilizador para píxeis de batata e cenoura juntos (b+c) para as 35 bandas de NDVI e as 6 bandas artificiais.

35 bandas de NDVI				6 bandas de parâmetros			
MV		RNA		MV		RNA	
EP	EU	EP	EU	EP	EU	EP	EU
98.89	90.52	41.89	62.01	69.57	91.95	84.23	77.32

Do ponto de vista do produtor e do utilizador pode concluir-se que, utilizando as 35 bandas de NDVI e recorrendo ao algoritmo MV, se obtêm os melhores resultados. Praticamente todos as parcelas de hortícolas correspondentes a batata e cenoura identificadas no terreno (98.9%) se encontram corretamente classificadas no mapa, enquanto que 90.5% das parcelas hortícolas representadas no resultado da classificação estão corretamente classificadas.

Os mapas de classificação produzidos a partir da implementação dos classificadores de MV e RNA, mediante as diferentes abordagens de bandas encontram-se ilustrados na Tabela 4.10.

Tabela 4.10 – Comparação visual dos resultados das classificações das 35 e 6 bandas (EG – Exatidão Global, IK – Índice Kappa, EP – Exatidão do Produtor e EU – Exatidão do Utilizador).

	Máxima Verosimilhança	Redes Neurais Artificiais
35 bandas	<p>EG:67.29% IK:61.62% EP(b+c):98.89% EU(b+c):90.52%</p>	<p>EG:21.90% IK:16.97% EP(b+c):41.89% EU(b+c):62.01%</p>
6 bandas	<p>EG:51.37% IK:44.88% EP(b+c):69.57% EU(b+c):91.95%</p>	<p>EG:39.62% IK:29.73% EP(b+c):84.23% EU(b+c):77.32%</p>

4.3.2. Estudo das RNA com máscara

A segunda análise consistiu na comparação dos resultados da classificação obtidos mediante alteração dos parâmetros do algoritmo RNA aplicadas às 6 bandas, mas com uma máscara. Com a delimitação da máscara consideraram-se essencialmente as zonas (píxeis) correspondentes às culturas a classificar, tendo por objetivo melhorar a exatidão dos resultados.

A Tabela 4.11 contém o número de polígonos que foram considerados como dados de treino e dados de validação para cada uma das classes definidas. Uma vez aplicada a máscara, o número de polígonos reduziu significativamente comparando com os valores listados na Tabela 4.5. A eliminação de polígonos evitou que píxeis NaN fossem confundidos com píxeis de culturas e garantiu um dos principais pré-requisitos para a classificação de dados em DR, onde objetos segmentados devem ter força descritiva e devem conter apenas píxeis de uma classe semântica de um mesmo grupo [Blaschke e Hay, 2000]. Os polígonos utilizados para parcelas de treino e de validação podem ser consultados no Anexo H.

Tabela 4.11 – Número de Parcelas de Treino (T) e de Validação (V) consideradas na comparação de Parâmetros para RNA.

	Máscara	Batata	Cenoura	Couve	Milho	Pastagem	Pousio	Prados	Tomate
T	5	12	8	1	6	5	9	13	2
V	6	12	8	1	6	3	8	19	2

4.3.2.1. Parametrização de modelos de RNA

A regra geral consistiu em alterar o valor de um parâmetro ao mesmo tempo que se mantinham os outros inalterados. Desta forma, o efeito de cada parâmetro pôde ser avaliado separadamente. Por exemplo, para investigar o impacto do número de camadas ocultas (*Hidden Layers*), três modelos de Redes Neurais Artificiais foram configurados com o número de camadas ocultas a variar entre 0 e 2, mantendo os restantes parâmetros inalterados. Existem outros parâmetros relevantes ao ajustamento da rede, tais como a taxa de treino (*Training Rate*), a taxa de momento de treino (*Training Momentum*), o critério de paragem (*Training RMS Exit Criteria*), o número de iterações (*Number of Training Iterations*) e o tamanho da contribuição do peso interno em relação ao nível de ativação do nó (*Training Threshold Contribution*), que permite ajustar as alterações no peso interno do nó. A taxa de treino determina a magnitude de ajustamento dos pesos. Uma taxa elevada acelera o treino, mas aumenta o risco de oscilações e a não convergência do resultado obtido. Uma taxa de momento de treino superior a zero permite uma alta taxa de treino sem oscilações. Um valor elevado deste parâmetro permite treinar a rede através de passos maiores quando comparado com um valor baixo. Para efetuar a paragem do treino da rede é utilizado um valor de erro quadrático mínimo. Se o erro quadrático for mais baixo que o critério de paragem estabelecido o treino finaliza, mesmo que o número de iterações definido não tenha sido atingido [Exelis, 2018].

Teste	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
TTC	0.9	0.9	0.9	0.9	0.9	0.2	0.5	0.7	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.7	0.8
TR	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.1	0.9	0.7	0.5	0.2	0.2	0.2	0.2	0.2	0.2	0.5	0.2
TM	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.1	0.5	0.9	0.9	0.9	0.9	0.9	0.9
TRMSE	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.05	0.08	0.1	0.1	0.1	0.1
NHL	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	2	1	1
NTI	1000	750	500	250	1250	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	500	900
EG	52.64	54.51	54.69	51.82	36.55	55.36	54.94	56.12	48.67		55.59		51.04	48.97	52.64		43.74	40.90	47.72	54.57
IK	41.46	44.81	45.50	41.70	25.32	44.56	40.74	46.07	36		45.72		39.15	36.33	41.46		27.28	21.93	37.94	44.95
EP (c + b)	82.69	82.69	82.66	82.69	82.68	84.57	83.36	82.77	82.69		82.68		82.69	82.69	82.69		91.60	86.27	82.68	82.67
EU (c + b)	87.66	88.99	89.01	89.02	88.99	82.26	83.79	87.03	84.60		89.01		84.62	84.78	87.66		81.16	65.49	88.99	89.01

(TTC - Training Threshold Contribution; TR - Training Rate; TM - Training Momentum; TRMSE - Training RMS Exit Criteria; NHL - Number of Hidden Layers; NTI - Number of Training Iterations)

Tabela 4.12 – Parâmetros considerados nas 20 configurações das RNA e respetivas estatísticas de classificação.

Na Tabela 4.12 encontram-se os parâmetros utilizados nas configurações do modelo e respetivos valores de exatidão global e índice *Kappa* obtidos para cada uma das combinações testadas. Fica claro que as exatidões da classificação podem variar para diferentes modelos neuronais associados a diferentes configurações de parâmetros internos.

- O ajuste dos pesos internos dos nós pode levar a melhores classificações, mas valores elevados também podem levar a más generalizações. Neste caso o valor do limiar não parece ser muito influente para a exatidão da classificação, no entanto a exatidão diminuiu quando o valor do limiar foi aumentado para 0.9;
- Uma taxa de treino mais alta acelera a execução, no entanto aumenta o risco de oscilações ou não-convergência do resultado de treino. Nenhuma relação óbvia entre a exatidão da classificação e a taxa de treino foi observada, uma vez que as exatidões se mantiveram iguais alterando os parâmetros;
- A exatidão da classificação foi superior para uma taxa de momento maior;
- Nenhuma relação óbvia entre a exatidão da classificação e o erro médio quadrático foi observada, uma vez que parâmetros diferentes obtiveram exatidões iguais;
- Aumentar o número de camadas ocultas não ajuda a melhorar a exatidão da classificação, pois os modelos com uma camada oculta mostraram melhores exatidões globais. Esta observação é consistente com os resultados obtidos por Zhou e Yang [2008];
- A exatidão da classificação aumentou quando o número de iterações aumentou, mas essa tendência diminuiu quando o número é bastante grande. Isto indica a ocorrência de excesso de treino, podendo reduzir a capacidade de generalização de redes neuronais. Um número de iterações mais elevado aumenta significativamente o tempo de execução.

De um modo geral as diferentes configurações acabaram por exibir valores não muito díspares uns dos outros em termos de exatidões, mas é possível detetar algumas diferenças na exatidão global e no índice *kappa* de cada classificação. Os Testes 8, 6, 10 e 3 foram os que apresentaram uma melhor exatidão global, de 56.12%, 55.59%, 55.36% e 54.69%, respetivamente. Os valores obtidos para o índice *kappa* correspondem a um desempenho moderado segundo o critério da Tabela 3.7. Estas classificações têm em comum o facto de quase todas apresentarem um erro de omissão de 100% para a classe correspondente à cultura de couve. Uma vez que não foi possível garantir um elevado número de parcelas de treino, assume-se que os frágeis resultados advenham do reduzido número de parcelas, o que contribui fortemente para a degradação da exatidão global e do índice *Kappa*.

Tabela 4.13 – Melhores resultados estatísticos para couve, cenoura e batata aquando comparação de parâmetros em RNA.

Matriz de Confusão (%)												
	Couve				Cenoura				Batata			
	T8	T6	T10	T3	T8	T6	T10	T3	T8	T6	T10	T3
Couve	0	0	0	0	0	0	0	0	0	0	0	0
Cenoura	0	0	0	12.5	45.43	45.56	48.99	70.5	2.57	2.61	8.43	23.85
Batata	91.67	100	91.67	79.17	29.33	31.31	25.73	4.22	84.98	86.56	78.99	63.56

Exatidão da Classificação (%)																
	Exatidão Produtor				Erro Omissão				Exatidão Utilizador				Erro Comissão			
	T8	T6	T10	T3	T8	T6	T10	T3	T8	T6	T10	T3	T8	T6	T10	T3
Couve	0	0	0	0	100	100	100	100	0	0	0	0	0	0	0	0
Cenoura	45.43	45.56	48.99	70.50	54.57	54.44	51.01	29.5	82.32	82.12	74.48	58.75	17.68	17.88	25.52	41.25
Batata	84.98	86.56	78.99	63.56	15.02	13.44	21.01	36.44	71.44	66.04	72.42	82.92	28.56	33.96	27.58	17.08

Na Tabela 4.13 são apresentadas as estatísticas alusivas às culturas em estudo, segundo os melhores resultados das classificações. Mais uma vez se verifica a total omissão das parcelas de couve para todas as combinações, o que leva a uma degradação da exatidão global. Genericamente existe elevada confusão entre as parcelas de batata e cenoura. Quando a exatidão do produtor é mais elevada para uma das culturas, será mais baixa para a outra, e vice-versa, ainda assim foram registadas exatidões do produtor na ordem dos 86.56% para a batata no Teste 6 e de 70.50% para a cenoura no Teste 3.

Tendo mais uma vez em conta as semelhanças que o ciclo vegetativo da cenoura e da batata apresentam, analisaram-se as exatidões do produtor e do utilizador das duas culturas como se fossem uma só classe.

Tabela 4.14 – Exatidões do Produtor e do Utilizador para píxeis de batata e cenoura juntos (b+c) para 6 bandas com máscara.





6 bandas de parâmetros artificiais com máscara							
Teste 8		Teste 6		Teste 10		Teste 3	
EP	EU	EP	EU	EP	EU	EP	EU
82.77	87.03	84.57	82.26	82.68	89.01	82.66	89.01

Comparando a Tabela 4.6 com a Tabela 4.12 podemos assumir que o uso da máscara e a otimização da rede neuronal levaram a um significativo melhoramento dos resultados, nomeadamente no que concerne à exatidão global aquando o uso das RNA, com um aumento de 16.5% no caso do Teste 8. Analisando também as Tabelas 4.9 e 4.14, podemos constatar que os Testes 3 e 10 permitiram um melhoramento de cerca de 12% na exatidão do utilizador.

Podemos ainda concluir que o Teste 10 é aquele que apresenta melhores resultados do ponto de vista do produtor e do utilizador, dado que a maior parte das parcelas hortícolas está corretamente identificada no mapa, com um baixo erro de comissão de 11% e que a maioria das parcelas hortícolas correspondentes a batata e cenoura identificadas no terreno (82.6%) encontram-se efetivamente bem classificadas no mapa, de onde 89.01% das parcelas hortícolas representadas no resultado da classificação estão corretamente classificadas. Uma vez que o resultado final mais relevante e apresentado aos beneficiários é o mapa classificado, assume-se a exatidão do utilizador como o indicativo mais importante.

Os mapas de classificação produzidos a partir da implementação do classificador RNA, mediante as diferentes abordagens de alteração de parâmetros encontram-se ilustrados na Tabela 4.15.

Tabela 4.15 – Comparação visual dos resultados das classificações com diferentes parâmetros de entrada nas RNA (EG – Exatidão Global, IK – Índice *Kappa*, EP – Exatidão do Produtor e EU – Exatidão do Utilizador).

Redes Neurais Artificiais	
	<p>Teste 8</p> <p>EG:56.12% IK:46.07% EP(b+c):82.77% EU(b+c):87.03%</p>
	<p>Teste 6</p> <p>EG:55.36% IK:44.56% EP(b+c):84.57% EU(b+c):82.26%</p>
	<p>Teste 10</p> <p>EG:55.59% IK:45.72% EP(b+c):82.68% EU(b+c):89.01%</p>
	<p>Teste 3</p> <p>EG:54.69% IK:45.50% EP(b+c):82.66% EU(b+c):89.01%</p>
<p>1: Thatata 2: Tcenoura 3: Tcouve 4: Tmascara 5: Tmilho 6: Tpastagem 7: Tpousio 8: Tprados 9: Ttomate</p>	

CAPÍTULO 5

CONCLUSÕES E PERSPETIVAS FUTURAS

5.1. PRINCIPAIS CONCLUSÕES

Ao longo deste trabalho de projeto foram aplicadas metodologias de Detecção Remota na extração de informação de imagens de alta resolução espacial e temporal (10 metros e 5 dias, respetivamente), para a caracterização fenológica dos vários tipos de culturas em estudo e respetiva classificação. Para esse efeito, foi avaliada a potencialidade de imagens da missão Sentinel-2 na geração de uma série temporal de bandas com o índice vegetativo NDVI, bem como na categorização de parâmetros que definem a curva do ciclo fenológico de cada cultura. Foi avaliado qual o algoritmo (Máxima Verosimilhança e Redes Neurais Artificiais) e que combinação de bandas (35 bandas de NDVI, 6 bandas de parâmetros e 6 bandas de parâmetros com máscara) permitiam produzir melhores resultados na classificação.

Foram recolhidas imagens de satélite que abrangiam o ciclo fenológico das culturas de forma a ser possível estudar cada uma através do seu comportamento, visível mediante a resposta espectral das imagens. Através do cálculo e análise da evolução temporal do NDVI foi possível identificar as fases fenológicas das culturas e verificar que as mesmas têm comportamentos semelhantes segundo este índice, no entanto, constata-se que a cultura de cenoura tende a apresentar um ciclo mais longo e uma maior amplitude, seguindo-se da de batata e da de couve, já mais breve.

O desenvolvimento de um algoritmo que calculava o valor mínimo, máximo, médio e desvio padrão de NDVI, bem como a amplitude e o número de dias correspondentes à curva do ciclo fenológico, píxel a píxel ao longo da série temporal (posição x , y , *datas*), permitiu gerar as 6 bandas artificiais de parâmetros. A definição da amplitude e do número de dias é independente de datas de sementeira e colheita mais precoces ou tardias, comparando entre diferentes parcelas de culturas iguais, o que pode ser visto com uma vantagem em termos de padronização de parâmetros para cada cultura.

A classificação das culturas através de diferentes combinações de informação e através de diferentes métodos de classificação permitiu retirar algumas conclusões. A classificação através do método da MV apresentou melhores resultados que o método das RNA na classificação multitemporal das 35 bandas de NDVI. Obteve-se uma exatidão global de 67.29%, de onde, do ponto de vista do produtor 98.89% das parcelas de batata e cenoura (juntas) foram classificadas corretamente. O facto de a cultura de couve apresentar apenas duas parcelas não permitiu retirar conclusões face ao seu comportamento. Esta situação realça que, para a obtenção de bons resultados, é necessária uma quantidade relativamente elevada de informação.

Para a classificação das 6 bandas de parâmetros artificiais, o classificador RNA teve um melhor desempenho que a MV, ainda que apenas tenha sido considerado razoável. Obteve-se uma exatidão global 39.62%, de onde, do ponto de vista do produtor 84.23% das parcelas de batata e cenoura foram corretamente classificadas. Com objetivo de estudar o comportamento das Redes Neurais Artificiais, procedeu-se à alteração dos parâmetros internos da rede, otimizando a rede e estudando o modo como estes influenciavam as exatidões. Dos 20 testes, o Teste 8 foi o que apresentou melhores resultados em termos de exatidão global, de 56.12%. Em termos de exatidão do produtor e do utilizador, o Teste 10 foi o que apresentou melhores resultados, de onde, do ponto de vista do utilizador 82.68% das parcelas

hortícolas foram classificadas corretamente, enquanto que 89.01% das parcelas hortícolas representadas no resultado da classificação estão corretamente classificadas.

Neste estudo, avaliou-se o impacto dos vários parâmetros internos das redes neuronais, treinados pelo método de retropropagação sobre o desempenho da classificação a partir de dados de Deteção Remota. Para otimizar o desempenho da classificação, recomenda-se de uma pequena taxa de aprendizagem, um momento moderado e um número de iterações também moderado. Espera-se que o estudo ajude a promover o uso operacional de Redes Neuronais Artificiais na classificação da cobertura hortícola.

Os resultados obtidos mostram-se satisfatórios, ainda que sujeitos a melhoria, comprovam a viabilidade e justificam a utilização da Deteção Remota e de procedimentos automáticos para a produção de cartografia. Importa referir que a metodologia aqui apresentada não se restringe apenas à região da Península de Setúbal, podendo ser aplicada a outras áreas de estudo. O trabalho de projeto desenvolvido pretende, por um lado demonstrar a aplicabilidade da metodologia desenvolvida e, por outro lado, ser um contributo para o desenvolvimento de novos estudos na área da Deteção Remota.

5.2. PERSPETIVAS FUTURAS

Em termos de perspetivas futuras, esta metodologia pode ser alvo de melhorias significativas no que concerne ao aumento do número de dados de treino e de validação, bem como à melhoria do algoritmo aquando a discriminação dos parâmetros, como o número de dias, por exemplo. Adicionalmente podem ainda ser geradas mais bandas de parâmetros, que se considerem relevantes na caracterização das curvas do ciclo fenológico de cada cultura. Com estas melhorias em vista, em vez de se ter uma série temporal ao longo de um ano, poder-se-ia usar uma série temporal ao longo de 3 ou 5 anos, já que uma alta resolução temporal permite comparações interanuais consistentes, apanhando períodos mais secos ou mais húmidos e identificando a variabilidade de NDVI, ou de outro índice vegetativo.

REFERÊNCIAS BIBLIOGRÁFICAS

Acharya, M. C., & Thapa, R. B. (2015) “Remote Sensing and Its Application In Agricultural Pest Management.” *The Journal of Agriculture and Environment*, Vol. 16, pp. 43–61.

Aziz, N., Minallah, N., Junaid, A., & Gul, K. (2017) “Performance Analysis of Artificial Neural Network Based Land Cover Classification.” *International Journal of Marine and Environmental Sciences*, Vol.11, No.5, pp. 422–426.

Blaschke, T., & Hay, G. J. (2000) “Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation.” *STROBL, J. et Al. (Hrsg.): Angewandte Geographische Informationsverarbeitung XII. Beiträge Zum AGIT-Symposium Salzburg*, Vol. 5, No. 3, pp. 12–23.

Bradter, U.; Thom, T. J.; Altringham, J. D.; Kunin, W. E.; Benton, T. G. (2011) “Prediction of national vegetation classification communities in the British uplands using multiple spatial scales, aerial images and the classifier random forest.” *Journal of Applied Ecology, Oxford*, Vol.48, No.4, pp. 1057-1065.

CCRS (2018): “Tutorial: Fundamentals of Remote Sensing”, <http://www.nrcan.gc.ca/earthsciences/geomatics/satellite-imagery-air-photos/satellite-imagery-products/educational-resources/9309>, acessado em junho de 2018.

Clark, M. L.; Aide, T. M.; Grau, H. R.; Riner, G. (2010) “A scalable approach to mapping annual land cover at 250 m using MODIS time series data: a case study in the Dry Chaco ecoregion of South America.” *Remote Sensing of Environment, New York*, Vol. 114, No. 11, pp. 2816-2832.

Congalton, R.G. & Green K. (2009) “Assessing the accuracy of remotely sensed data: principles and practices.” CRC Press, Second Edition. Boca Raton, United States of America.

Roy, D. P., Wulder, M. A., Loveland, T. R., C.E., W., Allen, R. G., Anderson, M. C., ... Zhu, Z. (2014) “Landsat-8: Science and product vision for terrestrial global change research.” *Remote Sensing of Environment*, Vol.145, pp.154–172.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Bargellini, P. (2012) “Sentinel-2: ESA’s Optical High-Resolution Mission for GMES Operational Services.” *Remote Sensing of Environment*, Vol.120, pp. 25–36.

Dusseux, P., Corpetti, T., Hubert-Moy, L., & Corgne, S. (2014) “Combined use of multi-temporal optical and Radar satellite images for grassland monitoring.” *Remote Sensing*, Vol.6, No.7, pp. 6163–6182.

Eklundh, L., Sjöström, M., Ardö, J., & Jönsson, P. (2012) “High resolution mapping of vegetation dynamics from Sentinel-2.” *Proceedings of the First Sentinel-2 Preparatory Symposium*, Vol. 707, No.1 23–27.

ESA (2018). <https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/product-types>, acessado em agosto de 2018.

- Exelis (2018). <https://www.harrisgeospatial.com/docs/NeuralNet.html>, acessado em junho de 2018.
- FAO (2009) “How to Feed the World in 2050 - Insights from an Expert Meeting at FAO - 24-26 June 2009”, *Population and Development Review* 35, pp.837–839.
- Fonseca, A. D. e Fernandes J.C. (2004): “Detecção Remota”, LIDEL – Edições Técnicas, Lisboa, Portugal.
- Fraser R.H., Olthof I., Carrière M., Deschamps A., Pouliot D. (2011) “Detecting long-term changes to vegetation in northern Canada using the Landsat satellite image archive.”, *Environmental Research Letters*, Vol. 6, No.4, pp.45-52.
- Haykin, S. (1994) “Neural Networks: A comprehensive foundation” Macmillan College Publishing Company: New York, NY, USA.
- He, L., Kong, F., & Shen, Z. (2006) “Artificial Neural Network Ensemble for Land Cover.” *Artificial Intelligence*, 10054–10057.
- Hebbbar, R., Ravishankar, H. M., Trivedi, S., Subramoniam, S. R., Raj, U., & Dadhwal, V. K. (2014) “Object oriented classification of high resolution data for inventory of horticultural crops.” *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, Vol. 40, No.8, pp. 745–749.
- Heumann B.W., Seaquist J.W., Eklundh L., Jönsson P. (2007) “AVHRR Derived Phenological Change in the Sahel and Soudan, Africa, 1982 – 2005”, *Remote Sensing of Environment*, Vol. 108, pp.385-392.
- Huang, J., Wang, H., Dai, Q., & Han, D. (2014) “Analysis of NDVI data for crop identification and yield estimation.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol.7, No.11, pp.4374–4384.
- IFAP (2018). http://www.ifap.minagricultura.pt/portal/page/portal/ifap_publico/GC_oifap#.WVNHvvyvIU, acessado em junho de 2018.
- IPMA (2018). <http://www.ipma.pt/pt/oclima/normais.clima/1981-2010/019/>, acessado em agosto de 2018.
- Jönsson A.M., Eklundh L., Hellström M., Barring L., Jönsson P. (2010) “Annual changes in MODIS vegetation indices of Swedish coniferous forests in relation to snow dynamics and tree phenology.” *Remote Sensing of Environment*, Vol. 114, No.11, pp.2719-2730.
- Karjalainen, M.; Kaartinen, H.; Hyypä, J. (2008) “Agricultural Monitoring Using Envisat Alternating Polarization SAR images.” *Photogrammetric Engineering & Remote Sensing*, Vol. 74, No.1, pp. 117-126.
- Karkee, M., Steward, B. L., Tang, L., & Aziz, S. A. (2009) “Quantifying sub-pixel signature of paddy rice field using an artificial neural network.” *Computers and Electronics in Agriculture*, Vol.65, No.1, pp. 65–76.
- Landis, J.R., Koch G.G. (1977) “The Measurement of Observer Agreement for Categorical Data.” *Biometrics*, Vol.33, No.1, pp. 159-174.

- Long, J., Lawrence, R., Greenwood, M. C., Marshall, L., & Miller, P. R. (2013) "Object-oriented crop classification using multitemporal ETM+ SLC-off imagery and random forest." *GIScience & Remote Sensing*, Vol.50, No.4, pp. 418–436.
- Masek J.G., Huang C., Wolfe R., Cohen W., Hall F., Kutler J., Nelson P. (2008) "North American forest disturbance mapped from a decadal Landsat record." *Remote Sensing of Environment*, Vol. 112, pp. 2914-2926.
- Menzel A. (2002) "Phenology: its importance to the global change community." *Climatic Change*, Vol. 54, pp. 379-385.
- Min, M., Ehsani, R., & Salyani, M. (2008) "Dynamic accuracy of GPS receivers in citrus orchards." *Applied Engineering in Agriculture*, Vol. 24, No. 6, pp. 861-868.
- Mishra, V. N., Prasad, R., Kumar, P., Gupta, D. K., & Srivastava, P. K. (2017) "Dual-polarimetric C band SAR data for land use/land cover classification by incorporating textural information." *Environmental Earth Sciences*, Vol.76, No. 1, article 26.
- Mondal, P., & Basu, M. (2009) "Adoption of precision agriculture technologies in India and in some developing countries: Scope, present status and strategies." *Progress in Natural Science*, Vol. 19, No. 6, pp.659–666.
- Neto, M. C., Pinto, P. A., e Coelho, J. P. P. (2005) *Tecnologias de Informação e Comunicação e a Agricultura*. Texto não publicado, Universidade de Lisboa: Sociedade Portuguesa de Inovação, Lisboa.
- Nijhawan, R., Roorkee, T., Srivastava, I., & Shukla, P, (2017). "Land Cover Classification Using Supervised and Unsupervised Learning Techniques", 2017 International Conference on Computational Intelligence in Data Science (ICCIDS), pp. 1–6.
- Otukei, J.R., Blaschke T. (2010) "Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms." *International Journal of Applied Earth Observation and Geoinformation*, Vol. 12, pp. S27-S31.
- Petropoulos, G. P., Vadrevu, K. P., Xanthopoulos, G., Karantounias, G., & Scholze, M. (2010) "A comparison of spectral angle mapper and artificial neural network classifiers combined with landsat TM imagery analysis for obtaining burnt area mapping." *Sensors*, Vol.10, No.3, pp.1967–1985.
- Piao S., Fang J., Zhou L., Guo Q., Henderson M., Ji W., Li Y., Tao S. (2003) "Interannual variations of monthly and seasonal normalized difference vegetation index (NDVI) in China from 1982 to 1999", *Journal of Geophysical Research*, Vol. 108, No. D14, 4401.
- Prata, C. S. R. (2010) *Palmela: Chão que dá uvas – A terra e o trabalho das gentes (1945 a 1958)*. Texto não publicado, Faculdade de Letras da Universidade de Lisboa (Departamento de História), Lisboa.
- Richards, J.A. & X. Jia (2006) *Remote Sensing Digital Image Analysis, An Introduction*. 4th edition, Springer, Berlin, Germany.
- Rodrigues, G., Pinto, C., Benites, F., & Melo, D. (2009) "Seleção para duração do ciclo vegetativo em batata e relação com a produtividade de tubérculos." *Horticultura Brasileira*, Vol. 27, pp. 280–285.

- Rouse, J. W., Hass, R. H., Schell, J. A., & Deering, D. W. (1973) "Monitoring vegetation systems in the great plains with ERTS." *Third Earth Resources Technology Satellite (ERTS) Symposium*, Vol. 1, pp. 309–317.
- Singh, A. (1989) "Digital Change Detection Techniques Using Remotely-Sensed Data." *International Journal of Remote Sensing*, Vol. 10, pp. 989-1003.
- Tagesson T., Mastepanov M., Tamstorf M.P., Eklundh L., Schubert P., Ekberg A., Sigsgaard C., Christensen T.R., Ström L. (2012) "Satellites reveal an increase in gross primary production in a greenlandic high arctic fen 1992–2008." *International journal of Applied Earth Observation and Geoinformation*, Vol.7, pp. 1101-1129.
- Usha, K., & Singh, B. (2013) "Potential applications of remote sensing in horticulture-A review." *Scientia Horticulturae*, Vol. 153, pp. 71–83.
- Warren, G., & Metternicht, G. (2005) "Agricultural Applications of High-Resolution Digital Multispectral Imagery." *Photogramm Eng Remote Sensing*, Vol. 71, No.5, pp. 595-602.
- Yang, C., Everitt, J. H., & Murden, D. (2011) "Evaluating high resolution SPOT 5 satellite imagery for crop identification." *Computers and Electronics in Agriculture*, Vol.75, No. 2, pp. 347–354.
- Yuan, H., Van Der Wiele, C. F., & Khorram, S. (2009) "An automated artificial neural network system for land use/land cover classification from landsat TM imagery." *Remote Sensing*, Vol. 1, No. 3, pp. 243–265.
- Zhou, L., & Yang, X. (2008) "Use of Neural Networks for Land Cover Classification from Remotely Sensed Imagery." *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXVII(Part B7), pp. 575–578.
- Zillmann, E., & Weichelt, H. (2014). Crop identification by means of seasonal statistics of RapidEye time series. 2014 The 3rd International Conference on Agro-Geoinformatics, Agro-Geoinformatics 2014.

ANEXOS

A. Datas de sementeira e colheita das Parcelas seleccionadas para o período de junho a dezembro de 2017, fornecidas pela PrimoHorto

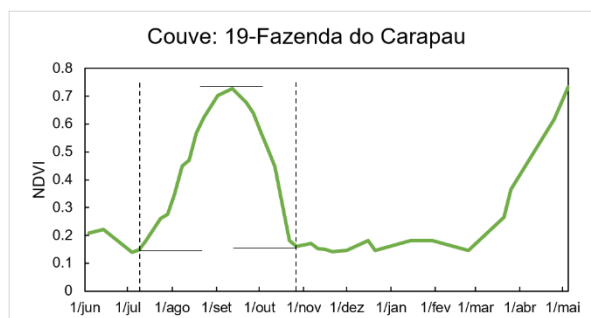
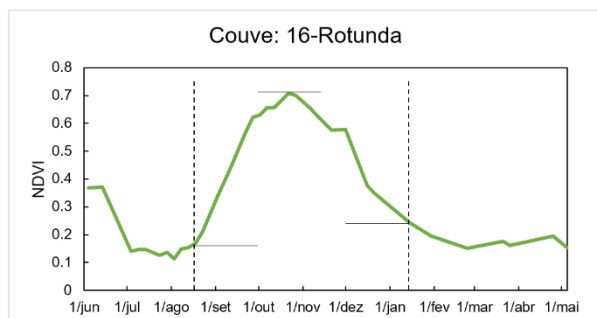
Produtor	Parcela	Parcela_Id	Cultura	Dat_Semen	Dat_Colhei
001	Herdade da Brejoeira	01B – 01	Batata	10/08/2017	.../12/2017
	Herdade da Brejoeira	02B – 02	Batata	10/08/2017	.../12/2017
	Urtiga Brava	07A – 07A	Cenoura	18/08/2017	17/01/2018
	Urtiga Brava	07B – 07B	Cenoura	18/08/2017	17/01/2018
	Urtiga Brava	8 – 08	Cenoura	25/06/2017	23/10/2017
	Urtiga Brava	9 – 09	Cenoura	31/07/2017	14/12/2017
	Urtiga Brava	10 – 10	Cenoura	31/07/2017	18/12/2017
	Urtiga Brava	11 – 11	Batata	24/08/2017	21/11/2017
	Quinta Velha	12 – 12	Batata	24/08/2017	01/02/2018
	Quinta Velha	13 – 13	Batata	24/08/2017	01/02/2018
003	Pegões	1 – 01	Cenoura	05/07/2017	23/10/2017
	Pegões	2 – 02	Cenoura	05/07/2017	23/10/2017
007	Rilvas_Brooke	2 – 02	Cenoura	10/06/2017	20/10/2017
	Palhavã	8 – 08	Batata	14/08/2017	24/02/2018
	Palhavã	10 – 10	Batata	14/08/2017	24/02/2018
	Malhadas	13 – 13	Batata	14/08/2017	15/01/2018
009	Rotunda	16 – 16	Couve	31/07/2017	02/02/2018
	Carapau	19 – 19	Couve	10/07/2017	10/12/2017
010	Camarate	21 – 21	Cenoura	30/08/2017	27/02/2018
	Camarate	20 – 20	Batata	25/08/2017	.../02/2018
011	Quinta das Flores	1 – 01	Cenoura	03/08/2017	26/12/2017
012	Quinta da Vinha	5 – 05	Batata	07/08/2017	.../11/2017
	Valmet	9A – 09A	Batata	07/08/2017	06/12/2017
	Valmet	9B – 09B	Cenoura	01/08/2017	13/12/2017
	Sela	13 – 13	Batata	.../08/2017	.../01/2018
	José Crespo	17 – 17	Batata	.../08/2017	.../01/2018
	São Francisco	1 – 01	Cenoura	05/07/2017	05/10/2017
013	São Francisco	2 – 02	Cenoura	05/07/2017	07/10/2017
	Quatro vales	08B – 08B	Cenoura	05/08/2017	08/12/2017
	Quatro vales	08D – 08D	Batata	10/08/2017	18/12/2017
	Coutadinha_LadyRoseta	1 – 01	Batata	.../08/2017	.../02/2018
014	Coutadinha_RedLady	1 – 01	Batata	.../08/2017	.../02/2018
	Quinta das Lebres	8 – 08A	Cenoura	.../07/2017	.../11/2017

	Quinta das Lebres	8 – 08B	Cenoura	.../07/2017	.../11/2017
017	Quinta do Passil	1C – 01C	Cenoura	16/08/2017	02/02/2018
	Quinta do Passil	1D – 01D	Batata	15/08/2017	.../01/2018
	Quinta do Passil	2A – 02A	Batata	15/08/2017	.../01/2018
	Quinta do Passil_LadyRosetta	2D – 02D	Cenoura	19/07/2017	02/02/2018
1028	Passarinhas	02A – 02A	Batata	20/08/2017	.../01/2018
	Passarinhas	02B – 02B	Cenoura	01/06/2017	09/10/2017

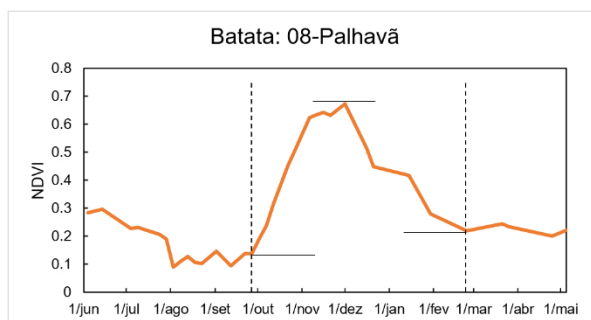
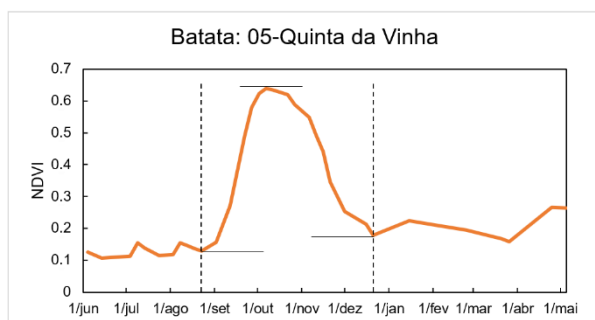
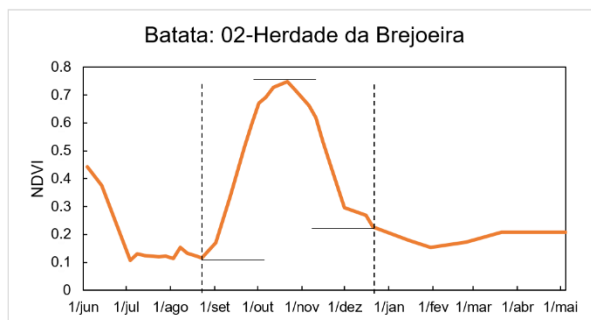
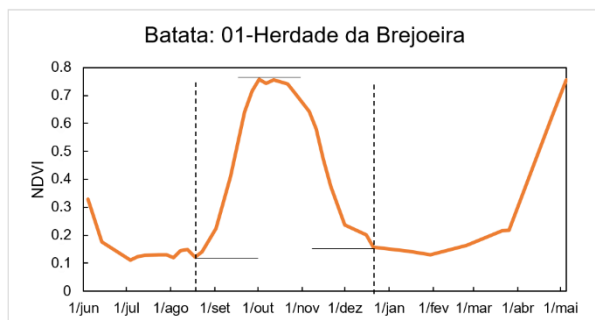
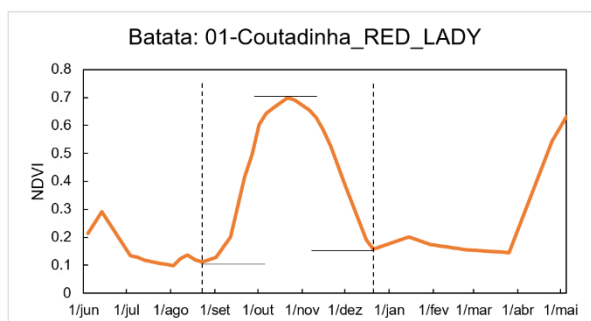
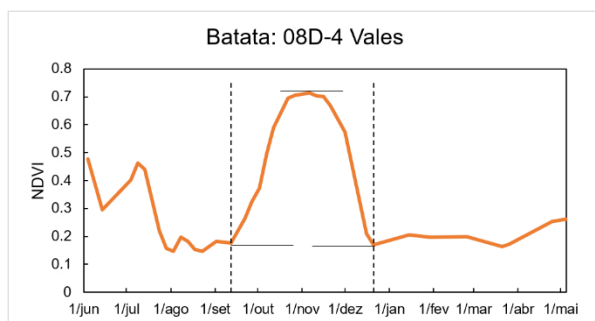
(*) Os campos assinalados a azul não possuem datas de sementeira e/ou colheita, apenas uma estimativa, dado que por vezes os agricultores não colocam o seu produto na central e a PrimoHorto não acompanha de perto a cultura. Sublinhado a cinzento encontram-se as parcelas removidas.

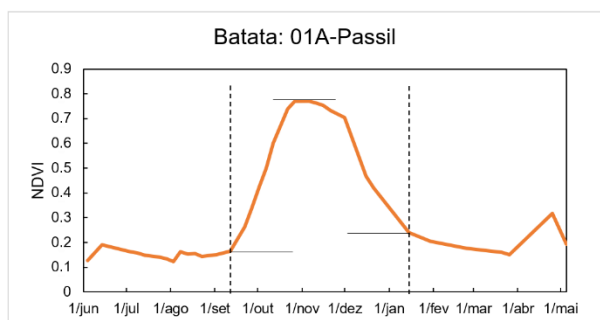
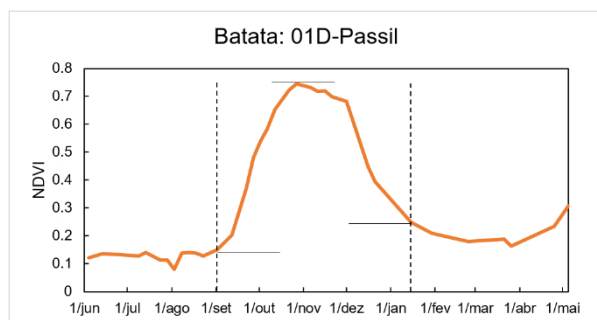
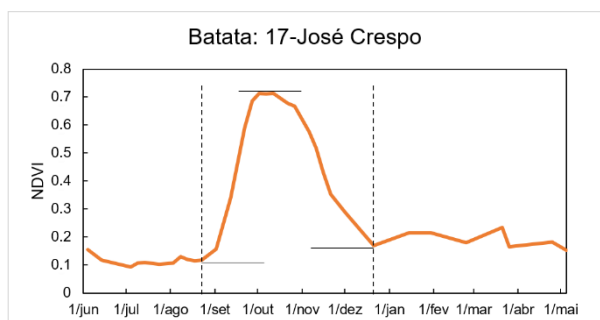
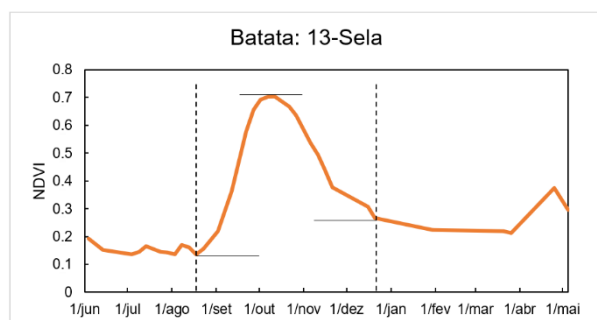
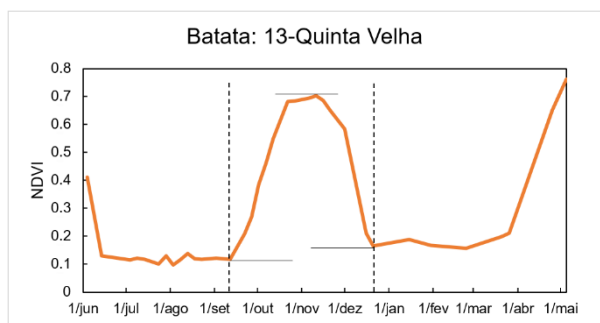
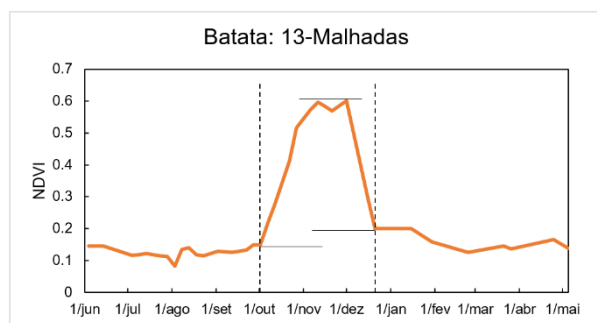
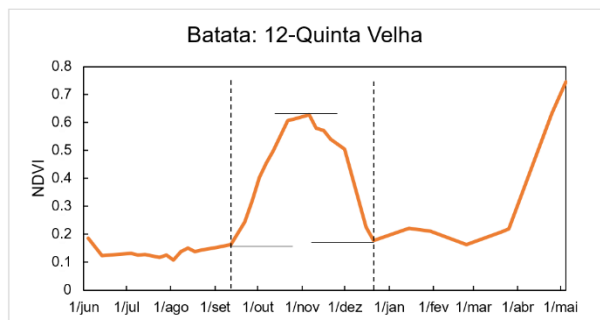
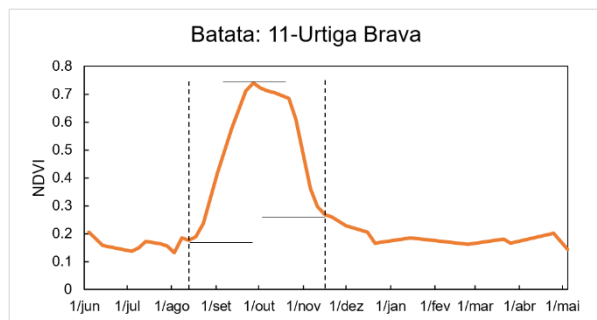
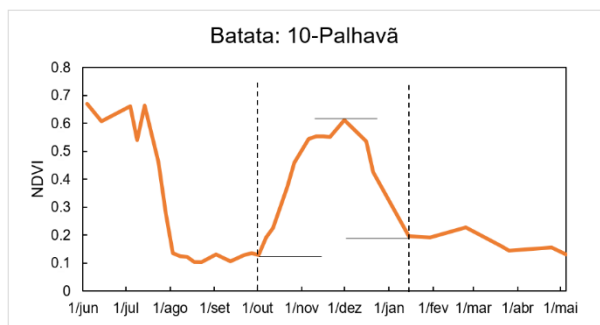
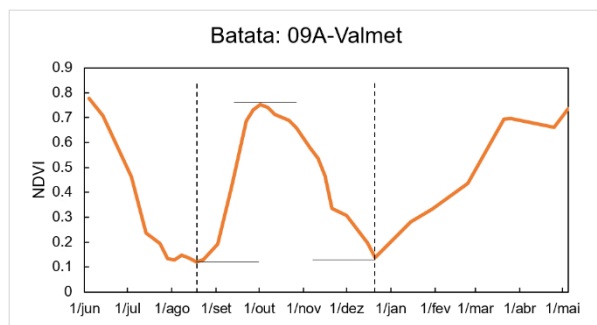
B. Evolução temporal do NDVI por parcelas (junho 2017 - maio 2018)

i. Parcelas de Couve

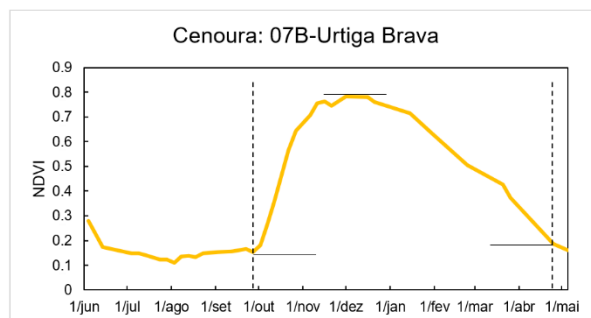
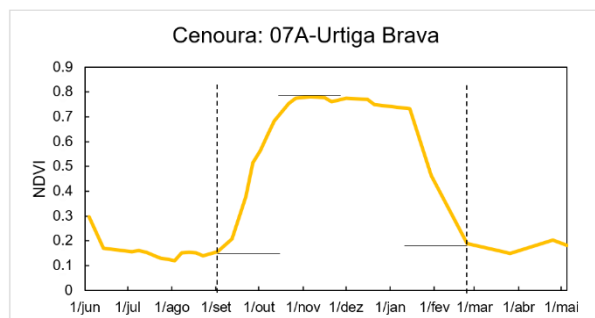
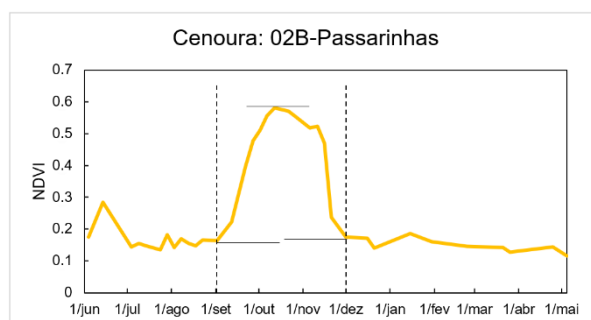
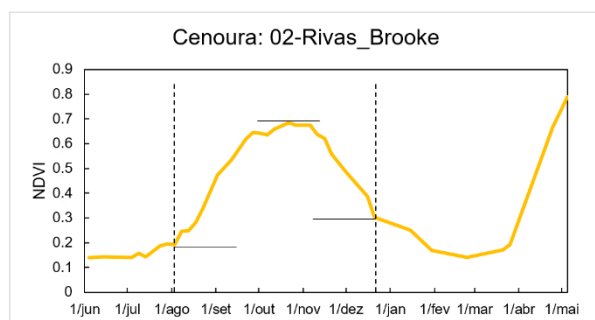
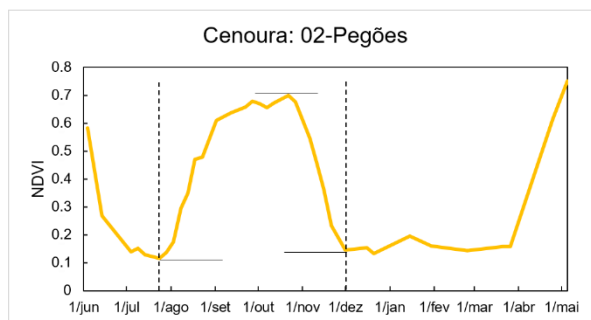
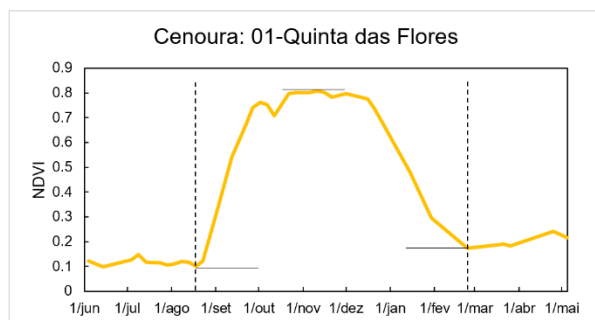
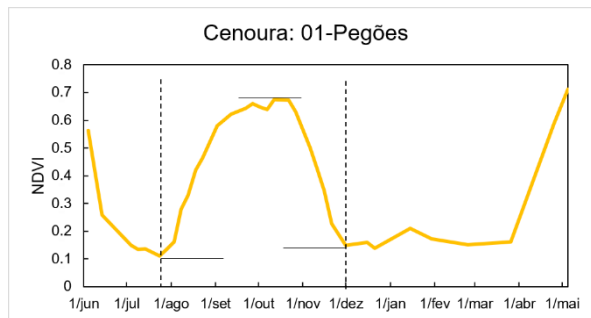
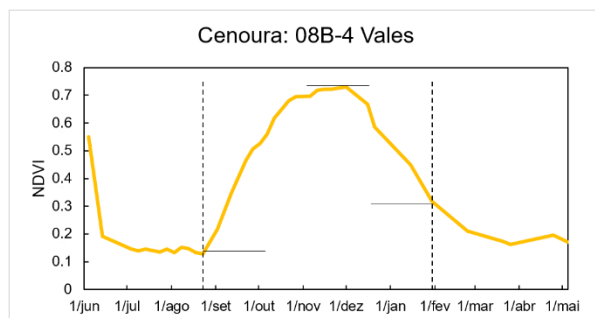
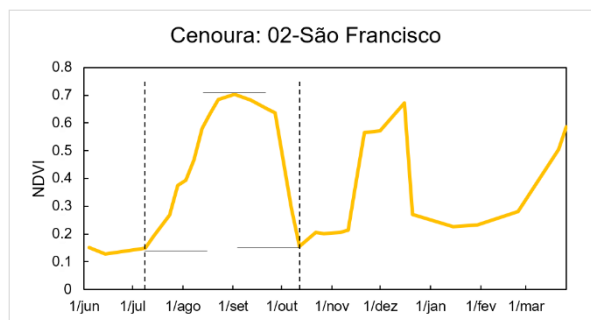
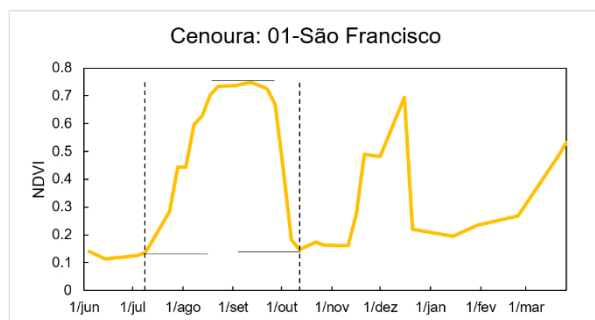


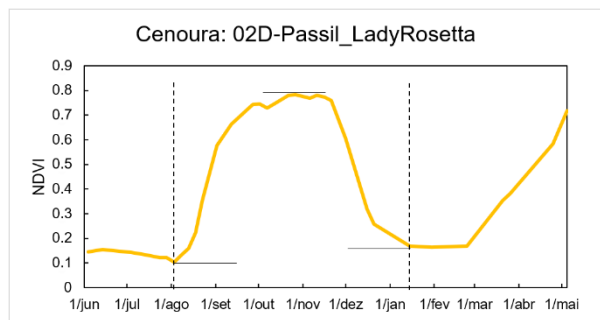
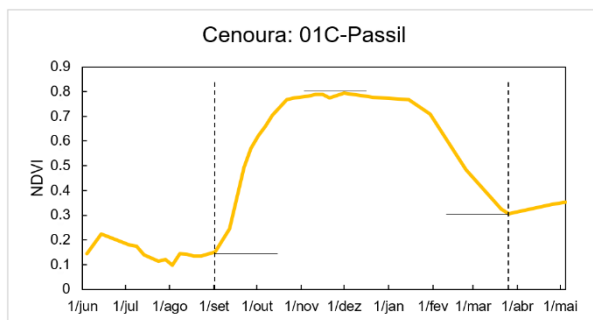
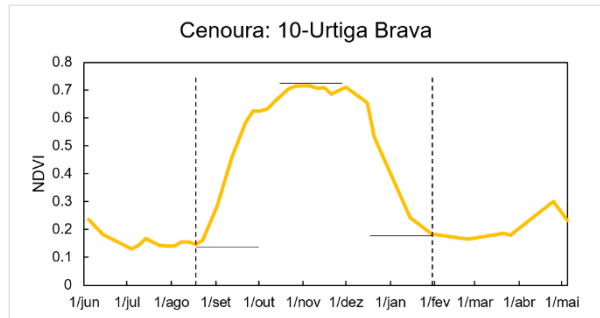
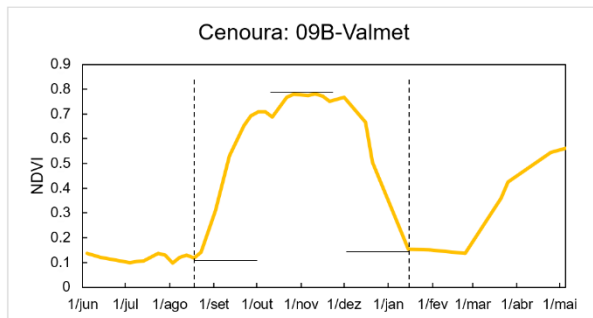
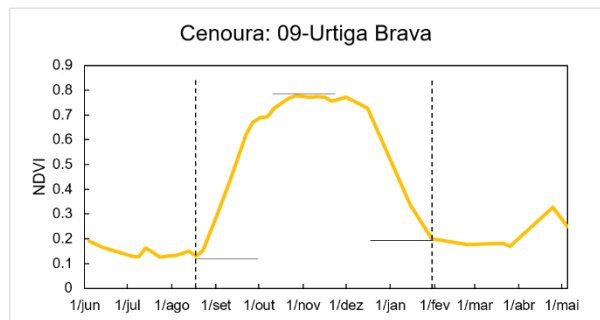
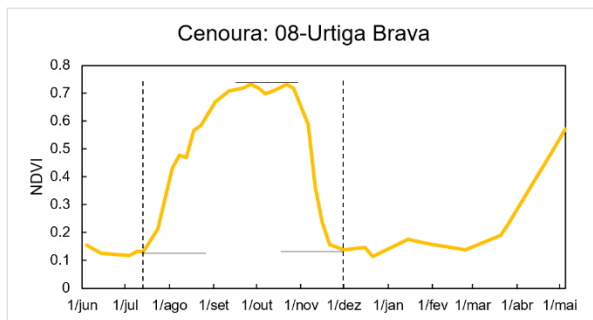
ii. Parcelas de Batata





iii. Parcelas de Cenoura





C. Valores médios por parcela calculados mediante análise visual em Excel**i. Parcelas de Couve**

Parcela	Max	Amp_{Esq}	Amp_{Dir}	Amp_{Média}	Nº dias
16-Rotunda	0.71	0.54	0.46	0.50	150
19-Fazenda do Carapau	0.73	0.58	0.57	0.57	110
Média	0.72	0.56	0.52	0.54	130

ii. Parcelas de Batata

Parcela	Max	Amp_{Esq}	Amp_{Dir}	Amp_{Média}	Nº dias
08D-4 Vales	0.71	0.54	0.54	0.54	150
01-Coutadinha_RED_LADY	0.70	0.59	0.54	0.56	120
01-Herdade da Brejoeira	0.75	0.64	0.60	0.62	125
02-Herdade da Brejoeira	0.75	0.63	0.52	0.58	120
05-Quinta da Vinha	0.64	0.51	0.46	0.48	120
08-Palhavã	0.67	0.53	0.45	0.49	150
09A-Valmet	0.75	0.62	0.61	0.62	120
10-Palhavã	0.67	0.54	0.47	0.51	105
11-Urtiga Brava	0.74	0.56	0.47	0.52	95
12-Quinta Velha	0.63	0.46	0.45	0.46	100
13-Malhadas	0.60	0.45	0.40	0.43	80
13-Quinta Velha	0.70	0.58	0.54	0.56	100
13-Sela	0.70	0.57	0.44	0.50	125
17-José Crespo	0.71	0.60	0.55	0.57	120
01D-Passil	0.76	0.60	0.50	0.55	135
02A-Passil	0.77	0.60	0.53	0.57	125
Média	0.70	0.56	0.51	0.54	118

iii. Parcelas de Cenoura

Parcela	Max	Amp_{Esq}	Amp_{Dir}	Amp_{Média}	Nº dias
01-São Francisco	0.75	0.61	0.60	0.61	95
02-São Francisco	0.70	0.55	0.55	0.55	95
08B-4 Vales	0.73	0.60	0.52	0.56	185
01-Pegões	0.67	0.56	0.53	0.54	130
01-Quinta das Flores	0.81	0.71	0.63	0.67	190
02-Pegões	0.70	0.59	0.55	0.57	130
02-Rivas_Brooke	0.68	0.49	0.43	0.46	165
02B-Passarinhas	0.58	0.42	0.41	0.41	90
07A-Urtiga Brava	0.78	0.57	0.59	0.58	165
07B-Urtiga Brava	0.78	0.63	0.59	0.61	210
08-Urtiga Brava	0.73	0.60	0.58	0.59	130
09-Urtiga Brava	0.78	0.65	0.58	0.61	165
09B- Valmet	0.78	0.66	0.63	0.64	150
10-Urtiga Brava	0.72	0.56	0.53	0.54	160
01C-Passil	0.79	0.64	0.49	0.57	205
02D-Passil_LadyRosetta	0.78	0.68	0.61	0.65	165
Média	0.74	0.60	0.55	0.57	152

D. Comparação do número de dias calculados entre as versões dos algoritmos

Cultura	Parcela	Referência Excel		Cálculo Matlab		
		Amp	Nº dias	Nº dias V. 1	Nº dias V. 2	Nº dias V. Half
Couve	16-Rotunda	0.50	150	134	194	121
	19-Fazenda do Carapau	0.57	110	119	119	118
Batata	08D-4 Vales	0.54	150	110	165	114
	01-Coutadinha RED LADY	0.56	120	118	158	118
	01-Herdade da Brejoeira	0.62	125	136	225	100
	02-Herdade da Brejoeira	0.58	120	126	253	117
	05-Quinta da Vinha	0.48	120	116	172	97
	08-Palhavã	0.49	150	160	267	122
	09A-Valmet	0.62	120	234	413	126
	10-Palhavã	0.51	105	199	220	320
	11-Urtiga Brava	0.52	95	111	147	92
	12-Quinta Velha	0.46	100	98	150	100
	13-Malhadas	0.43	80	124	156	102
	13-Quinta Velha	0.56	100	109	171	108
	13-Sela	0.50	125	133	217	116
	17 - José Crespo	0.57	120	127	179	98
	01D-Passil	0.55	135	179	254	130
	02A-Passil	0.57	125	181	289	121
	01-São Francisco	0.61	95	122	205	132
	02-São Francisco	0.55	95	96	217	124
	08B-4 Vales	0.56	185	198	253	163
	01-Pegões	0.54	130	125	252	156
Cenoura	01-Quinta das Flores	0.67	190	187	265	150
	02-Pegões	0.57	130	115	177	140
	02-Rilvas_Brooke	0.46	165	158	307	166
	02B-Passarinhas	0.41	90	108	134	99
	07A-Urtiga Brava	0.58	165	209	293	152
	07B-Urtiga Brava	0.61	210	365	411	136
	08-Urtiga Brava	0.59	130	134	244	161
	09-Urtiga Brava	0.61	165	198	271	157
	09B- Valmet	0.64	150	167	291	157
	10-Urtiga Brava	0.54	160	181	255	148
	01C-Passil	0.57	205	264	353	160
	02D-Passil_LadyRosetta	0.65	165	167	359	177

E. Valores médios por parcela calculados através da Versão 1 do algoritmo em *Matlab*

i. Parcelas de Couve

Parcela	Máximo	Mínimo	Amplitude	Número de Dias
19-Fazenda do Carapau	0.75	0.12	0.56	103
16-Rotunda	0.78	0.11	0.57	130
Média	0.77	0.11	0.56	117

ii. Parcelas de Batata

Parcela	Máximo	Mínimo	Amplitude	Número de Dias
01-Herdade da Brejoeira	0.68	0.17	0.43	528
02-Herdade da Brejoeira	0.78	0.11	0.63	134
05-Quinta da Vinha	0.76	0.11	0.59	126
08-Palhavã	0.68	0.10	0.51	118
11-Urtiga Brava	0.69	0.09	0.50	147
13-Quinta Velha	0.75	0.13	0.56	111
13-Sela	0.78	0.09	0.56	106
17-José Crespo	0.73	0.11	0.52	124
08D-4 Vales	0.74	0.14	0.57	108
01-Coutadinha_RED_LADY	0.73	0.10	0.58	117
09A-Valmet	0.79	0.12	0.61	149
10-Palhavã	0.70	0.10	0.54	194
12-Quinta Velha	0.75	0.10	0.49	91
13-Malhadas	0.63	0.08	0.48	121
01D-Passil	0.77	0.08	0.61	180
02A-Passil	0.79	0.11	0.63	179
Média	0.73	0.11	0.55	158

iii. Parcela de Cenoura

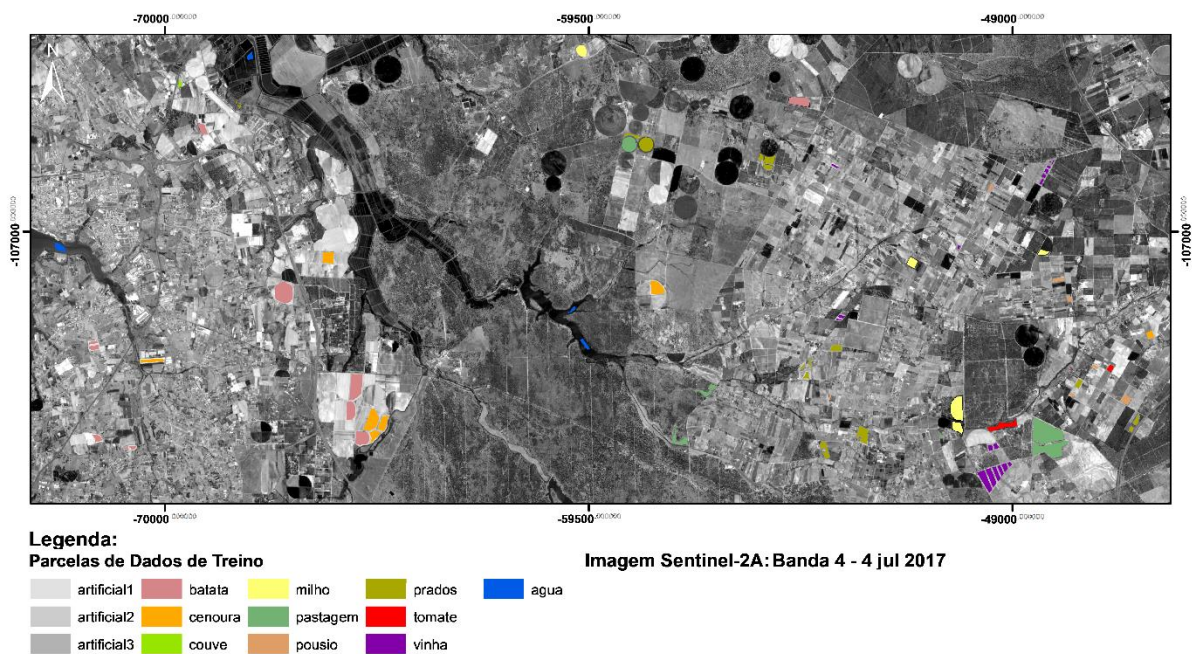
Parcela	Máximo	Mínimo	Amplitude	Número de Dias
02-São Francisco	0.78	0.12	0.60	94
08B-4 Vales	0.77	0.12	0.61	199
02-Pegões	0.79	0.11	0.57	110
07B-Urtiga Brava	0.79	0.11	0.65	363
09-Urtiga Brava	0.79	0.12	0.65	196
09B-Valmet	0.80	0.09	0.65	163
10-Urtiga Brava	0.75	0.13	0.58	178
02D-Passil_LadyRosetta	0.82	0.10	0.68	165
01-São Francisco	0.79	0.10	0.61	115
01-Pegões	0.74	0.11	0.57	125
01-Quinta das Flores	0.83	0.09	0.68	186
02-Rilvas_Brooke	0.81	0.12	0.55	153
02B-Passarinhas	0.62	0.10	0.46	101
07A-Urtiga Brava	0.80	0.12	0.65	206
08-Urtiga Brava	0.76	0.11	0.51	119
01C-Passil	0.83	0.10	0.64	241
Média	0.78	0.11	0.60	170

F. Máscara aplicada

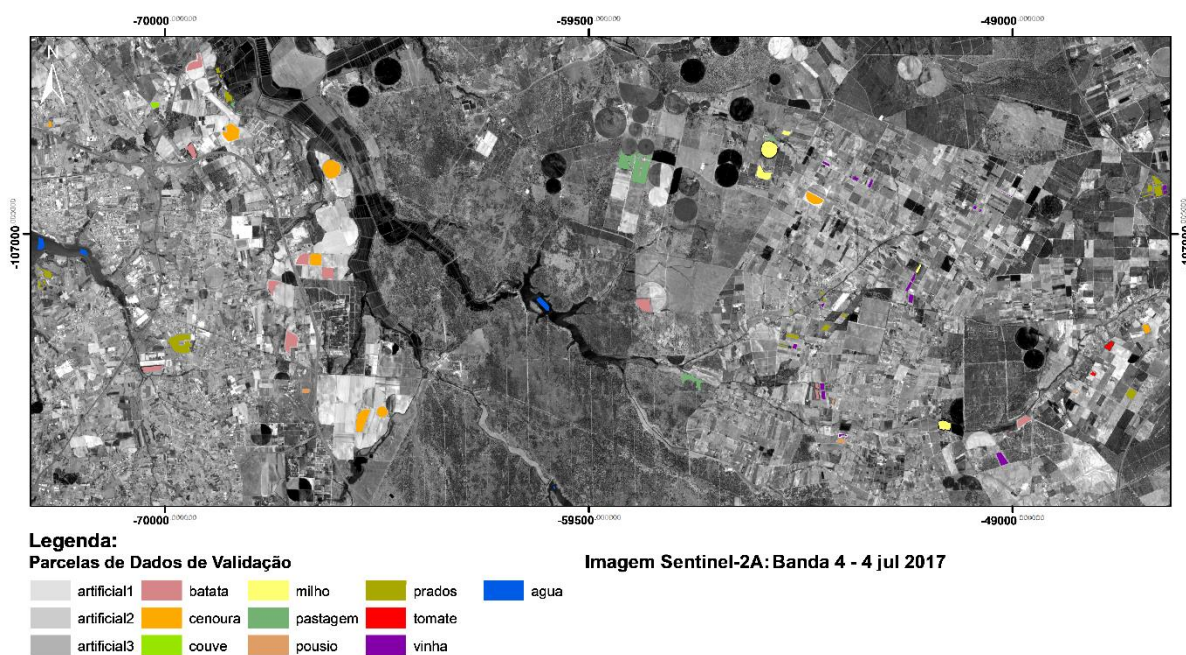


G. Dados de treino e de validação para a comparação de resultados na classificação das culturas usando MV e RNA

i. Dados de Treino

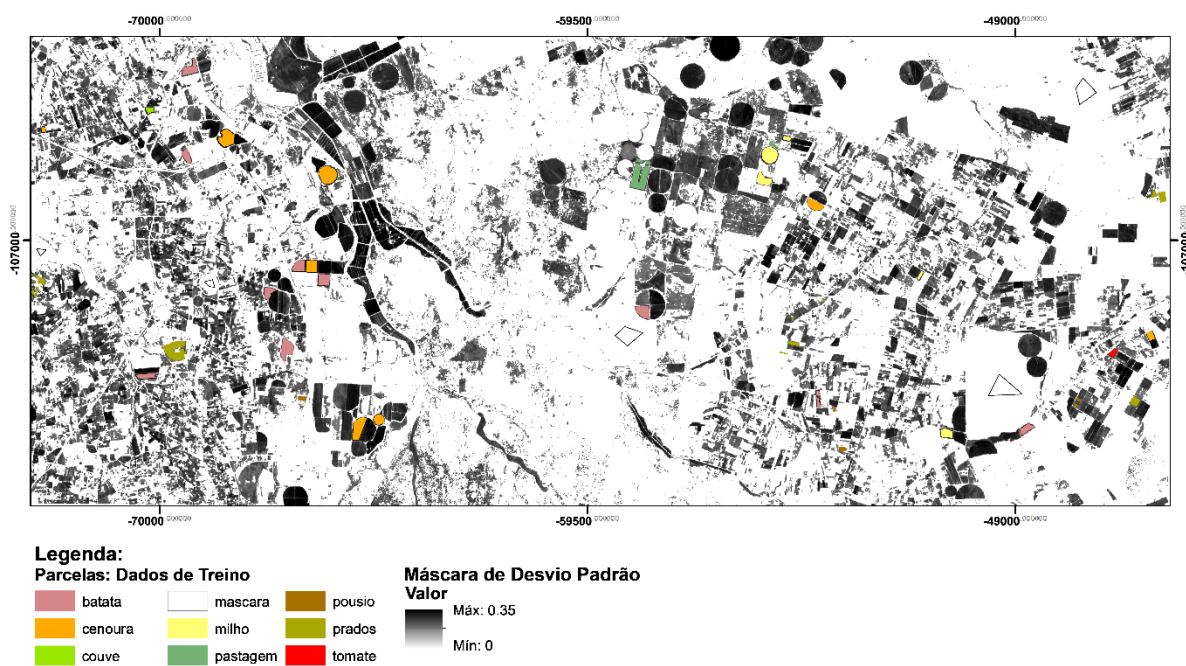


ii. Dados de Validação



H. Dados de treino e de validação para a comparação de resultados na classificação das culturas mediante os diferentes parâmetros das RNA

i. Dados de Treino



ii. Dados de Validação

